



**Universidade de Brasília**

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# **Detecção de Doença de Alzheimer a partir de Neuroimagens Utilizando Redes Neurais Convolucionais**

Maximillian Fan Xavier

Monografia apresentada como requisito parcial  
para conclusão do Curso de Engenharia da Computação

Orientador  
Prof. Dr. Alexandre Zaghetto

Brasília  
2018



**Universidade de Brasília**

**Instituto de Ciências Exatas  
Departamento de Ciência da Computação**

# **Deteccção de Doença de Alzheimer a partir de Neuroimagens Utilizando Redes Neurais Convolucionais**

Maximillian Fan Xavier

Monografia apresentada como requisito parcial  
para conclusão do Curso de Engenharia da Computação

Prof. Dr. Alexandre Zaghetto (Orientador)  
CIC/UnB

Prof. Me. Mateus Mendelson Esteves da Silva    Prof. Dr. Marcus Vinícius Chaffim Costa  
UnB Gama/FGA    UnB Gama/FGA

Dr. Diogo Neves Jacó  
Clínica Goiana de Diagnósticos

Prof. Dr. José Edil  
Coordenador do Curso de Engenharia da Computação

Brasília, 1 de dezembro de 2018

# Dedicatória

Aos meus pais, Claudia e Sergio, por toda criação que me concederam, sempre aconselhando e apoiando minhas escolhas.

Dedico também à minha avó Áurea, por, assim como meus pais, ter me criado e sempre buscado o melhor para mim.

À minha irmã Paola, por ser a amiga mais fiel.

E, por fim, dedico a todos que estiveram comigo nessa jornada, principalmente Deus, que iluminou minha trajetória.

# Agradecimentos

Agradeço profundamente a **Deus**, por toda saúde e perseverança que me concedeu.

Aos meus pais, **Claudia** e **Sergio**, por todo amor e carinho recebido até hoje.

À minha irmã **Paola**, por sempre me apoiar, em qualquer situação, sendo a mais bela das amigadas.

Aos meus tios **Silvana** e **Cleder**, por abrir um caminho grandioso para mim, o qual estou conquistando hoje.

Um agradecimento mais do que especial aos grandes amigos que formei nessa trajetória: **Matheus Veleci**, **Murilo Medeiros** e **Rafael Costa**. Só cheguei onde cheguei por contar com vocês em diversos momentos.

Ao meu orientador **Alexandre Zaghetto** por todo conhecimento compartilhado e atenção nos diversos momentos.

Agradeço também à *Alzheimer's Disease Neuroimaging Initiative* (ADNI) pelo acesso aos dados e qualidade/relevância da informação.

Um grande agradecimento a **Universidade de Brasília**, por ter me concedido a oportunidade de viver ótimos momentos, fazer grandes amigos e alcançar muitas conquistas.

À minha companheira de todas as horas e momentos **Miriã Carvalho**, por toda paciência, amor e carinho fornecidos em cada etapa. Obrigado mesmo, eu te amo.

Por fim, obrigado a todos e todas que estiveram comigo nessa jornada, família ou amigos, cada palavra ou abraço foi fundamental para essa conquista.

# Resumo

A doença de Alzheimer, doença neuro degenerativa, apresenta minúcias pontuais que precisam ser analisadas de forma precisa e detalhada, sendo um trabalho de grande custo temporal e técnico para os profissionais da saúde responsáveis por esse tipo de atividade. Logo, uma solução que automatize essa tarefa se mostra fundamental.

Para concretizarmos isso, nossa proposta é elaborar uma rede neural convolucional, capaz de classificar pacientes em diferentes estágios da doença a partir de imagens de ressonância magnética.

A rede convolucional realiza procedimentos de convolução e subamostragem em imagens digitais, extraindo características que são utilizadas como parâmetros de aprendizagem pela rede neural artificial, garantindo robustez e agilidade.

**Palavras-chave:** Alzheimer, redes neurais, processamento de imagens, redes convolucionais, aprendizagem de máquina

# Abstract

The Alzheimer disease is a neurodegenerative disease that affects millions of people. It consists in a progressive loss of neurons, which can be diagnosed with image analysis by a health professional.

The goal of this paper is to make a classifier based in convolutional neural networks, which can differentiate digital images provided by magnetic resonance.

The convolutional neural network aims to extract features from these digital images, executing convolutional operations and subsampling, making these features to be used as parameters to neural network, with robustness and agility.

**Keywords:** Alzheimer, neural networks, image processing, convolutional networks, machine learning

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Revisão Bibliográfica</b>	<b>3</b>
2.1	A Doença de Alzheimer . . . . .	3
2.2	Imagens Digitais . . . . .	6
2.3	Filtragem e Correção . . . . .	8
2.4	Segmentação . . . . .	9
2.5	Geração de imagens do sistema nervoso . . . . .	12
2.6	O arquivo NIFTI . . . . .	12
2.7	Inteligência Artificial . . . . .	13
2.7.1	Aprendizagem de Máquina . . . . .	14
2.7.2	Redes Neurais Convolucionais . . . . .	17
2.8	Transferência de Aprendizagem . . . . .	25
2.9	Trabalhos Correlatos . . . . .	27
<b>3</b>	<b>Solução Proposta</b>	<b>33</b>
3.1	Base de dados . . . . .	33
3.2	Pré processamento das imagens . . . . .	34
3.3	Extração de características via transferência de aprendizagem . . . . .	35
3.4	Classificação da imagem . . . . .	36
<b>4</b>	<b>Resultados Experimentais</b>	<b>41</b>
4.1	Elaboração e treinamento de rede neural convolucional . . . . .	42
4.2	Resultados . . . . .	43
4.2.1	Primeiro ambiente: conjunto com todas imagens . . . . .	45
4.2.2	Segundo ambiente: conjunto reduzido . . . . .	53
4.2.3	Terceiro ambiente: conjunto reduzido com realce de bordas . . . . .	61
<b>5</b>	<b>Conclusão</b>	<b>69</b>





# Lista de Figuras

2.1	Sistema nervoso central. . . . .	3
2.2	Estruturas afetadas diretamente e indiretamente pela doença de Alzheimer - (a) córtex cerebral; (b) hipocampo; e (c) ventrículos. . . . .	5
2.3	Regiões afetadas - no lado esquerdo um cérebro saudável e no lado direito um cérebro afetado pela doença. Em vermelho a região cortical; em verde os ventrículos e em azul o hipocampo. . . . .	6
2.4	Imagens digitais - (a) imagem colorida; e (b) imagem em níveis de cinza. .	7
2.5	Quantização de imagens digitais em níveis de cinza - (a) pixels de 8 bits e 256 níveis de cinza; (b) pixels de 4 bits e 16 níveis de cinza; (c) pixels de 2 bits e 4 níveis de cinza; e (d) pixels de 1 bit e 2 níveis de cinza. . . . .	7
2.6	Aplicação do filtro de mediana - (a) região da imagem original; e (b) região após aplicação do filtro. . . . .	8
2.7	Filtro de mediana aplicado na MRI - (a) imagem original; e (b) imagem resultante. . . . .	9
2.8	<i>Power-law</i> aplicado na MRI - (a) imagem original; e (b) imagem resultante.	10
2.9	Segmentação de imagem utilizando método Canny - (a) imagem em níveis de cinza; e (b) imagem segmentada. . . . .	11
2.10	Técnica para geração de imagens digitais do corpo humano a partir da utilização de ressonância magnética para avaliação de estágios da doença de Alzheimer. . . . .	12
2.11	Eixos do corpo humano. . . . .	13
2.12	Estrutura de um neurônio biológico. . . . .	18
2.13	Estrutura de um neurônio artificial. . . . .	19
2.14	Comportamento das funções de ativação - (a) função de limiar; (b) função sigmoide com $a = 5$ e $a = 10$ ; (c) função tangente hiperbólica com $x = 3$ ; e (d) função ReLU. . . . .	21
2.15	Arquitetura de camada única com alimentação adiante. . . . .	22
2.16	Arquitetura de múltiplas camadas com alimentação adiante. . . . .	22

2.17	Ilustração da operação de convolução - (a) primeira etapa da convolução da imagem com o filtro; (b) segunda etapa da convolução da imagem com o filtro; (c) décima terceira etapa da convolução da imagem com o filtro; e (d) última etapa da convolução da imagem com o filtro. . . . .	24
2.18	Operação de agregação. . . . .	25
2.19	Arquitetura geral de uma rede neural convolucional. . . . .	25
2.20	Impacto na performance de treinamento com e sem utilização de transferência de aprendizagem. Imagem adaptada de [1]. . . . .	26
3.1	Fatias do volume tri dimensional. . . . .	34
3.2	Imagem após aplicação de filtro de mediana, transformação <i>power-law</i> e determinação de ROI. . . . .	35
3.3	Imagem após aplicação de filtro de mediana, transformação <i>power-law</i> , determinação de ROI e segmentação com método <i>Canny</i> . . . . .	36
3.4	Evolução das redes neurais convolucionais profundas. Imagem adaptada de [2]. . . . .	37
3.5	Exemplo de comportamento da função de entropia categórica cruzada. . . .	38
3.6	Estrutura geral da rede neural convolucional. . . . .	40
4.1	Exemplos de imagens utilizadas como entrada para a rede convolucional - (a) portador da doença de Alzheimer; (b) paciente com distúrbios cognitivos leves; e (c) paciente saudável. . . . .	42
4.2	Classificação multi classe no primeiro ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	45
4.2	Classificação multi classe no primeiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. . . . .	46
4.3	Classificação binária (Alzheimer x Saudáveis) no primeiro ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	47
4.3	Classificação binária (Alzheimer x Saudáveis) no primeiro ambiente de experimentação - (b) Gráfico de perda x iterações ; e (c) Matriz de confusão. . . . .	48
4.4	Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no primeiro ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	49
4.4	Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no primeiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. . . . .	50
4.5	Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no primeiro ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	51

4.5	Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no primeiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. . . . .	52
4.6	Classificação multi classe no segundo ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	53
4.6	Classificação multi classe no segundo ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. . . . .	54
4.7	Classificação binária (Alzheimer x Saudáveis) no segundo ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	55
4.7	Classificação binária (Alzheimer x Saudáveis) no segundo ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. .	56
4.8	Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no segundo ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	57
4.8	Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no segundo ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. . . . .	58
4.9	Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no segundo ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	59
4.9	Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no segundo ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. . . . .	60
4.10	Classificação multi classe no terceiro ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	61
4.10	Classificação multi classe no terceiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. . . . .	62
4.11	Classificação binária (Alzheimer x Saudáveis) no terceiro ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	63
4.11	Classificação binária (Alzheimer x Saudáveis) no terceiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. .	64
4.12	Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no terceiro ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	65
4.12	Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no terceiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. . . . .	66
4.13	Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no terceiro ambiente de experimentação - (a) Gráfico de acurácia x iterações; . . . . .	67

4.13 Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no terceiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão. . . . .	68
---	----

# Lista de Tabelas

4.1 Resultados experimentais. . . . .	44
---------------------------------------	----

# Lista de Abreviaturas e Siglas

**ADNI** *Alzheimer’s Disease Neuroimaging Initiative.*

**AI** *Artificial Intelligence.*

**ConvNets** *Convolutional Networks.*

**GMM** *Gaussian Mixture Models.*

**GPU** *Graphics Processing Unit.*

**ILSVRC** *ImageNet Large-Scale Visual Recognition Challenge.*

**MCI** *Mild Cognitive Impairment.*

**MRI** *Magnetic Resonance Imaging.*

**NIFTI** *Neuroimaging Informatics Technology Initiative.*

**PET** *Positron emission tomography.*

**PLS** *Partial Least Squares.*

**PNG** *Portable Network Graphics.*

**RAM** *Random Access Memory.*

**RELU** *Rectified Linear Unit.*

**ResNet** *Residual Network.*

**RGB** *Red Green Blue.*

**ROI** *Region of Interest.*

**SGD** *Stochastic Gradient Descent.*

**SVM** *Support Vector Machine.*

**VGG16** *Visual Geometry Group - 16 Layers.*

# Capítulo 1

## Introdução

O diagnóstico de inúmeras doenças passa pela análise minuciosa de características específicas de um paciente, sejam elas clínicas (exame de sangue, pressão arterial, etc.) ou demográficas (idade, atuação profissional, cidade em que vive, etc.). Essa análise é realizada por um profissional da área da saúde com determinada especialização, que as vezes não conta com instrumentos de grande auxílio ou não dispõe de experiência diante de um quadro clínico até então desconhecido.

Dentre as diversas doenças já estudadas pelo homem, a doença de Alzheimer continua tomando grande atenção. A doença de Alzheimer é uma doença neuro degenerativa que afeta milhões de pessoas, ocasionando em perda gradativa das células cerebrais (neurônios), que implicam em problemas motores e cognitivos [3]. Tanto suas origens quanto seu tratamento ainda não apresentam embasamento teórico definitivo, havendo muitas pesquisas nesse campo para ratificar esses aspectos.

O diagnóstico por sua vez já é alcançável, através da análise progressiva de exames de sangue, condições motoras e cognitivas e principalmente pela observação de imagens médicas. As imagens médicas são obtidas na maioria dos casos através de ressonância magnética, que propiciam uma visualização moderada das regiões cerebrais. O profissional da saúde, após analisar criteriosamente as imagens, bem como demais dados clínicos, pode indicar a existência ou não dessa patologia.

Porém, a análise minuciosa dessas imagens passa pela expertise do profissional bem como pela qualidade do equipamento que gera essas imagens. Logo, o profissional precisa, além da experiência, dispôr de imagens que não apresentem borrões ou defeitos que impeçam o diagnóstico com precisão. Portanto, se faz necessária uma ferramenta que auxilie esse profissional na tomada de decisão no momento do diagnóstico e que apresente elevada confiabilidade.

Com esse objetivo, este trabalho visa providenciar uma ferramenta computacional capaz de realizar a classificação de imagens médicas de pacientes em diferentes estágios



da doença, quando presente.

Através de procedimentos de *deep learning*, utilizamos redes neurais convolucionais que extraem características minuciosas de imagens digitais advindas de ressonância magnética e classificam-nas em um determinado estágio da doença de Alzheimer. Obviamente, esse método consiste em uma ferramenta que visa auxiliar a tomada de decisão do profissional de saúde, uma vez que dados clínicos corroboram para um diagnóstico mais robusto.

No Capítulo 2 apresentamos os conceitos teóricos empreendidos nesse trabalho. Dentre eles, estão aspectos sobre a doença de Alzheimer, processamento de imagens digitais, inteligência artificial e redes neurais convolucionais.

O Capítulo 3 exhibe a metodologia proposta por esse trabalho, com intuito de explicitar ao leitor as ferramentas utilizadas e a arquitetura projetada.

Os resultados do trabalho em questão são apresentados no Capítulo 4, com objetivo de exhibir cenários, parâmetros utilizados e performance obtida.

Por fim, no Capítulo 5 realiza-se novamente uma breve análise do problema que culminou nesse trabalho, bem como possíveis tarefas para uma continuação futura.

# Capítulo 2

## Revisão Bibliográfica

O objetivo deste capítulo é apresentar alguns conceitos e estruturas referentes aos procedimentos realizados no Capítulo 3.

### 2.1 A Doença de Alzheimer

A doença de Alzheimer é uma doença neuro degenerativa que pode gerar complicações graves para um indivíduo. Para entendermos melhor de que forma essa doença pode afetar o portador, precisamos compreender melhor o sistema nervoso e as estruturas que sofrem com essa patologia.

O sistema nervoso de todos os mamíferos apresenta duas divisões: o sistema nervoso central e o sistema nervoso periférico [4]. O sistema nervoso central são as porções do sistema nervoso que estão envolvidas pelos ossos, sendo elas o encéfalo e a medula espinhal. O encéfalo localiza-se inteiramente no crânio e é constituído por cérebro, cerebelo e tronco encefálico. A Figura 2.1 exibe o sistema nervoso central.

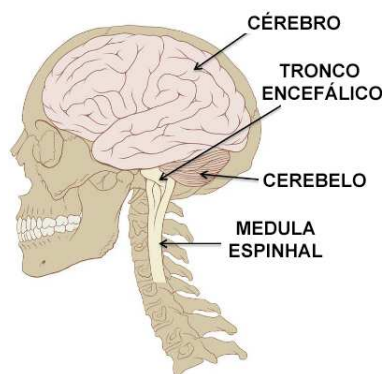


Figura 2.1: Sistema nervoso central<sup>1</sup>.

---

<sup>1</sup>Adaptada de: TodaMatéria - <https://bit.ly/2ARp1FV>. Acesso: 30/11/2018

O cérebro é a porção mais larga do encéfalo, composto por dois hemisférios cerebrais. Em geral, o hemisfério cerebral direito recebe sensações e controla o movimento do lado esquerdo do corpo; de forma semelhante, o hemisfério cerebral esquerdo está envolvido com sensações e movimentos do lado direito do corpo. O cerebelo, situado atrás do cérebro, é uma porção menor, que é primariamente um centro para controle do movimento que possui extensivas conexões com o cérebro e a medula espinhal. O tronco encefálico é a porção restante que é um conjunto complexo de fibras e células, que, em parte, serve para enviar informação do cérebro à medula espinhal e ao cerebelo e da medula espinhal e cerebelo ao cérebro. No entanto, o tronco encefálico é também uma região que regula funções vitais, como respiração, consciência e controla da temperatura. Por mais primitivo que seja, o tronco encefálico é uma estrutura essencial para a vida, pois danos ao cérebro e cerebelo são passíveis de sobrevivência, enquanto que danos ao tronco encefálico significam morte rápida.

A medula espinhal é envolvida pela coluna vertebral óssea e está colada ao tronco encefálico. Ela é responsável por conduzir informação da pele, articulações e músculos ao encéfalo e vice versa.

No encéfalo existem três estruturas que tomam maior atenção, devido à sua importância com a doença de Alzheimer: córtex, hipocampo e ventrículos. O córtex cerebral corresponde à camada mais externa do cérebro, composta de regiões e sulcos, sendo formado por massa cinzenta (neurônios, capilares e células de glia - células que auxiliam a operação dos neurônios). Essa estrutura desempenha um papel importantíssimo em funções complexas do cérebro como memória, atenção, consciência, linguagem e percepção. O hipocampo, estrutura pertencente ao sistema límbico (unidade responsável primariamente pelas emoções e comportamentos) é um componente essencial para o cérebro, pois trata da formação de memórias a longo prazo e de navegação espacial. Os ventrículos constituem o sistema ventricular, sendo quatro regiões cavernosas no encéfalo (uma em cada hemisfério do cérebro, uma entre os dois hemisférios e uma logo abaixo dessa), preenchidas pelo líquido cefalorraquidiano, também conhecido por fluído cerebrospinal (que age como um amortecedor para o córtex e medula espinhal, na “limpeza” do metabolismo de drogas ou substâncias que se difundem no cérebro através do sangue e na distribuição de substâncias para as células cerebrais), sendo sua função a produção desse líquido. Essas três estruturas são exibidas na Figura 2.2.

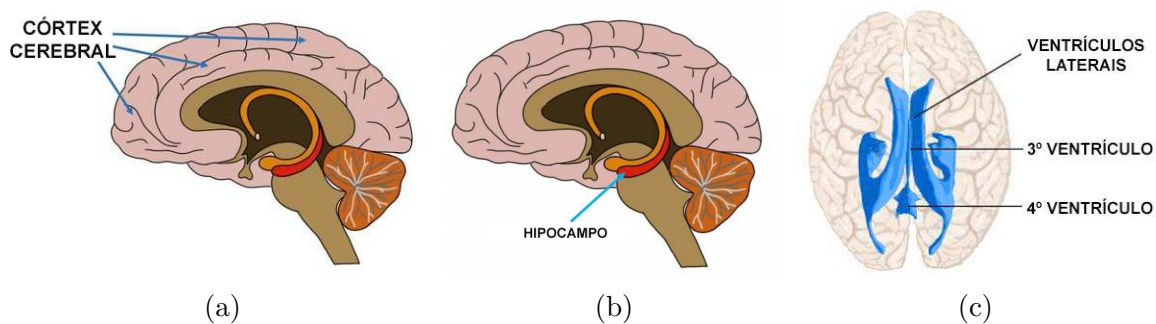


Figura 2.2: Estruturas afetadas diretamente e indiretamente pela doença de Alzheimer - (a) córtex cerebral<sup>2</sup>; (b) hipocampo<sup>3</sup>; e (c) ventrículos<sup>4</sup>.

Conhecendo um pouco mais da estrutura do sistema nervoso podemos então aprofundar o estudo sobre a doença de Alzheimer, ficando a cargo do leitor obter mais detalhes em [3]. A doença de Alzheimer é uma doença neuro degenerativa que afeta a massa cinzenta, caracterizada pela perda progressiva de neurônios e consequente manifestação de demência (alteração progressiva da função cognitiva). Essa perda neuronal ocorre de forma seletiva, afetando um grupo de células e deixando outros grupos intactos.

Sua ocorrência é mais comum em idosos, apresentando inicialmente prejuízos às vezes irrelevantes, como alteração de humor e comportamento. Posteriormente ocorre desorientação progressiva, perda de memória e afasia (capacidade de formulação e compreensão da linguagem), sintomas esses que já representam um quadro mais grave de disfunção cortical. Em um período de 5 a 10 anos o indivíduo tende a ficar totalmente incapacitado, mudo e imóvel.

A maioria dos pacientes são assintomáticos antes dos 50 anos de idade, porém a incidência aumenta com o passar do tempo. Os pacientes mais graves possuem idade em torno de 85 e 89 anos. Esse aumento progressivo da probabilidade de incidência acarreta em sérios problemas médicos, sociais e econômicos para países que possuem muitos idosos em sua composição populacional.

Do ponto de vista genético molecular, anomalia fundamental na doença de Alzheimer é o acúmulo de duas proteínas,  $A\beta$  e  $\tau$ , em regiões específicas do cérebro, provavelmente como resultado da produção excessiva. Essas duas proteínas em excesso, quando combinadas, influenciam na comunicação dos neurônios (alterando a morfologicamente os microtúbulos, dentro do axônio) bem como em seu funcionamento interno, levando-os à morte.

<sup>2</sup>Adaptada de: Neuroscientifically Challenged - <https://bit.ly/2praR9X>. Acesso: 30/11/2018.

<sup>3</sup>Adaptada de: Neuroscientifically Challenged - <https://bit.ly/2Ds3sjo>. Acesso: 30/11/2018.

<sup>4</sup>Adaptada de: Neuroanatomy - <https://bit.ly/2RvrTzt>. Acesso: 30/11/2018.

Analisando por uma visão macro, a doença de Alzheimer provoca diferentes graus de atrofia cortical, atrofia do hipocampo e considerável aumento dos sulcos cerebrais, com maior ênfase aos lobos frontais. Com as atrofias, os ventrículos tendem a compensar os espaços, ocorrência conhecida como hidrocefalia *ex vacuo* [3]. As implicações estruturais da doença são exibidas na Figura 2.3.

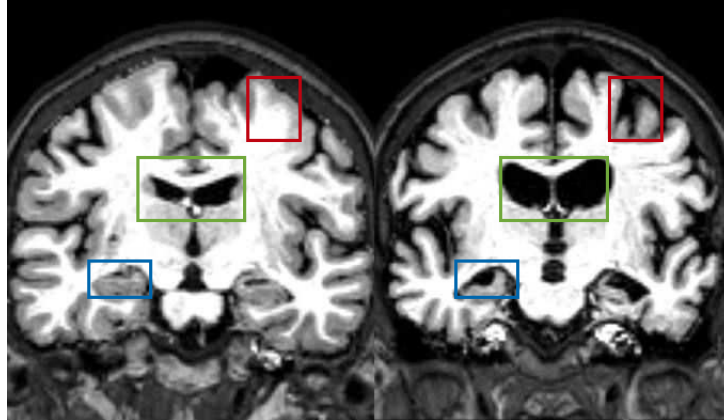


Figura 2.3: Regiões afetadas - no lado esquerdo um cérebro saudável e no lado direito um cérebro afetado pela doença. Em vermelho a região cortical; em verde os ventrículos e em azul o hipocampo<sup>5</sup>.

Para realizar um diagnóstico mais definitivo dessa doença um exame do tecido encefálico precisa ser executado. Entretanto, a combinação de elementos clínicos e métodos radiológicos modernos permitem diagnósticos com precisão que varia entre 80% e 90% dos casos.

## 2.2 Imagens Digitais

Imagens digitais podem ser representadas de forma colorida ou em níveis de cinza (conforme Figura 2.4), utilizando tonalidades de cor entre o branco e o preto. As imagens em escala de cinza são representadas por matrizes bidimensionais de  $M$  linhas e  $N$  colunas, sendo que cada elemento dessa matriz é conhecido como pixel (*picture element*), representado pela Equação 2.1; enquanto que imagens coloridas, por exemplo, no espaço de cor RGB (do inglês *Red, Green, Blue*) são imagens compostas por três matrizes bidimensionais de  $M$  linhas e  $N$  colunas, sendo cada uma das matrizes uma representação da componente vermelha (R), componente verde (G) e componente azul (B), respectivamente. Podemos relacionar uma imagem em níveis de cinza ( $Y$ ) com uma imagem colorida, no espaço de cor RGB, através da Equação 2.2.

---

<sup>5</sup>Adaptada de: Pinterest - <https://bit.ly/2G8WcLo>. Acesso: 30/11/2018.

$$f(x, y) = I, \quad (2.1)$$

onde  $x$  e  $y$  representam as coordenadas da matriz (linha e coluna) e  $I$  a intensidade de cor denotada por um valor inteiro.

$$Y = 0,299 \times R + 0,587 \times G + 0,114 \times B \quad (2.2)$$



Figura 2.4: Imagens digitais - (a) imagem colorida<sup>6</sup>; e (b) imagem em níveis de cinza<sup>7</sup>.

O nível de cinza de cada pixel varia em uma faixa de valores possíveis de acordo com a quantidade de bits disponíveis para representar esse elemento. Quanto maior a quantidade de bits disponíveis, maior será o detalhamento da imagem em sua representação e vice-versa, como visto na Figura 2.5. A Equação 2.3 exhibe a relação entre os níveis de cinza e os bits utilizados na representação da imagem.

$$L(b) = 2^b, \quad (2.3)$$

onde  $b$  indica a quantidade de bits disponíveis e  $L(b)$  a quantidade de tons na escala de cinza.



Figura 2.5: Quantização de imagens digitais em níveis de cinza - (a) pixels de 8 bits e 256 níveis de cinza; (b) pixels de 4 bits e 16 níveis de cinza; (c) pixels de 2 bits e 4 níveis de cinza; e (d) pixels de 1 bit e 2 níveis de cinza<sup>8</sup>.

<sup>6</sup>Disponível em: Wikipedia - <https://bit.ly/2QCq5HU>. Acesso: 30/11/2018.

<sup>7</sup>Disponível em: Digital Watermarking - <https://bit.ly/2rkCbXG>. Acesso: 30/11/2018.

<sup>8</sup>Adaptada de: Digital Watermarking - <https://bit.ly/2rkCbXG>. Acesso: 30/11/2018.

Como a metodologia deste trabalho utiliza apenas imagens em níveis de cinza assume-se seu uso deste ponto em diante.

## 2.3 Filtragem e Correção

As imagens digitais podem sofrer problemas em sua representação, seja por problemas no equipamento de captura (câmeras fotográficas digitais) ou no momento da aquisição da imagem, por exemplo, exposição intensa à luz em um determinado cenário. Para contornar esses obstáculos faz-se necessário o uso de alguns procedimentos que corrijam esses problemas, também conhecidos como ruídos, e para isso temos os procedimentos de filtragem e correção/ajuste de contraste e saturação da imagem.

A utilização de filtros em imagens digitais tem objetivos diversos, desde o efeito de borramento da imagem até o realce de bordas ou elementos específicos. Nesse trabalho utilizou-se o filtro de mediana, que aplica um efeito de borramento sobre a imagem, removendo pequenos detalhes desnecessários, porém sem perder tanta informação de borda dos elementos.

O filtro de mediana consiste em calcular o valor mediano de intensidade dos pixels em uma determinada região da imagem, geralmente uma região quadrada e de tamanho ímpar (3x3, 5x5, etc.), e substituir o pixel central dessa região com esse valor. Essa operação é realizada em toda a imagem, da esquerda para a direita e de cima para baixo, utilizando um parâmetro de movimento (*stride*), indicando a quantidade de pixels que a região irá avançar em cada etapa. A Figura 2.6 exibe a região de uma imagem com pixel central de valor 7, bem como a região resultante com pixel de valor 4. A Figura 2.7 exibe uma aplicação real do filtro de mediana.

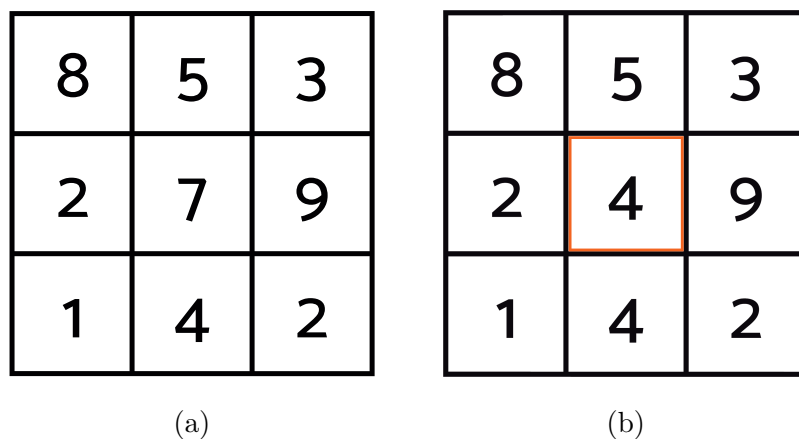


Figura 2.6: Aplicação do filtro de mediana - (a) região da imagem original; e (b) região após aplicação do filtro.

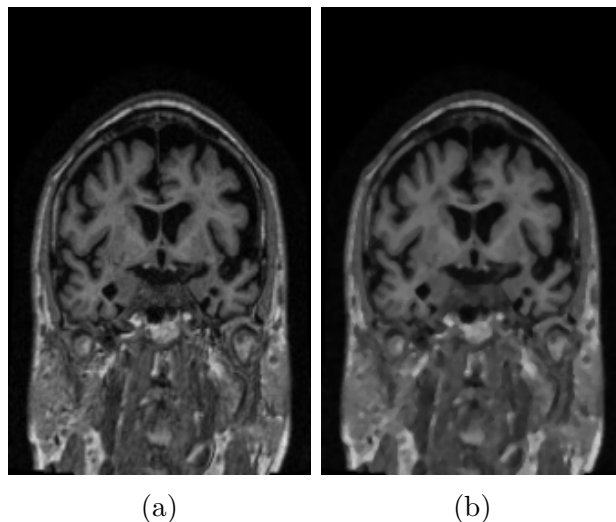


Figura 2.7: Filtro de mediana aplicado na MRI - (a) imagem original; e (b) imagem resultante.

Para ajustarmos o contraste e a saturação de imagens em níveis de cinza podemos utilizar métodos (transformações) que mapeiam um determinado nível de cinza para outro de acordo com alguma regra. Um desses métodos é o ajuste por escala exponencial de contraste (*power-law*). Esse método relaciona duas quantidades, onde uma mudança em um desses valores resulta em uma variação proporcional no outro valor. A Equação 2.4 exibe a fórmula convencional para aplicar esse método:

$$s = cr^{\gamma}, \quad (2.4)$$

onde  $c$  e  $\gamma$  são constantes,  $r$  é o pixel com um determinado nível de cinza e  $s$  é o nível de cinza resultante.

Muitos dispositivos utilizados para capturar, imprimir ou exibir imagens possuem algum conjunto de parâmetros pré estabelecidos para a *power-law*, o que pode ocasionar em imagens mais claras ou mais escuras do que a originalmente obtida. Para corrigirmos esse problema, aplicamos a correção do expoente  $\gamma$ , variando seu valor para baixos valores, por exemplo, valores entre  $[0,0.4]$  (quando a imagem está sendo exibida de forma demasiadamente escura) ou altos valores, por exemplo,  $[1.5,5.0]$  (quando a imagem está sendo exibida de forma demasiadamente clara). A Figura 2.8 exibe uma aplicação real da correção através da *power-law*.

## 2.4 Segmentação

A técnica de segmentação de imagens consiste em subdividir a imagem em suas regiões constituintes ou objetos. O grau de divisão que é realizado depende do problema a ser



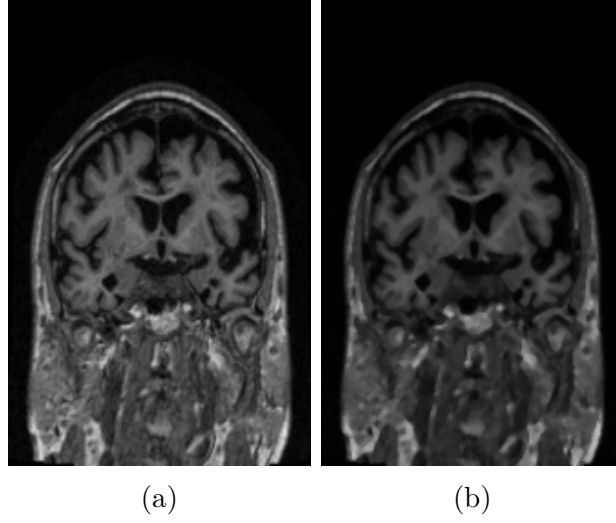


Figura 2.8: *Power-law* aplicado na MRI - (a) imagem original; e (b) imagem resultante.

resolvido.

Boa parte dos algoritmos são baseados em uma das duas propriedades básicas de valores de intensidade: descontinuidade e similaridade. Na primeira categoria, a abordagem consiste em dividir os elementos de acordo com as mudanças abruptas de intensidade entre os elementos, por exemplo, bordas. Já na segunda, o método consiste em dividir a imagem de acordo com algum conjunto de critérios predefinidos de similaridade entre os elementos.

Existem muitos métodos/algoritmos para segmentar uma imagem, desde a separação entre limiares até a presença de objetos de determinado formato. Na Figura 2.9 observamos a aplicação do método *Canny* de segmentação, que consiste no seguinte procedimento (podendo ser aprofundado em [5]):

1. Criar duas máscaras Gaussianas unidimensionais,  $G(x)$  e  $G(y)$  (expressas pelas Equações 2.5 a 2.6) para serem convoluídas com uma imagem  $I$ . As formas das máscaras Gaussianas dependem da escolha do desvio padrão,  $\sigma$ , utilizado;

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} \quad (2.5)$$

$$G(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} \quad (2.6)$$

2. Criar duas máscaras unidimensionais para o cálculo da primeira derivada das Gaussianas, nas direções  $x$  (linha) e  $y$  (coluna), denominadas por  $G'(x)$  e  $G'(y)$ , e expressas pelas Equações 2.7 a 2.8;

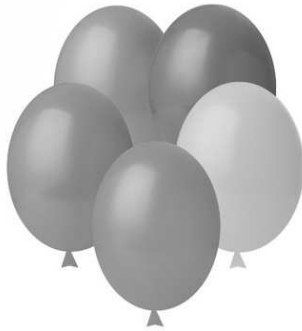
$$G'(x) = \frac{-x \frac{-x^2}{2\sigma^2}}{\sqrt{2\pi\sigma^3}} \quad (2.7)$$

$$G'(y) = \frac{-y \frac{-y^2}{2\sigma^2}}{\sqrt{2\pi\sigma^3}} \quad (2.8)$$

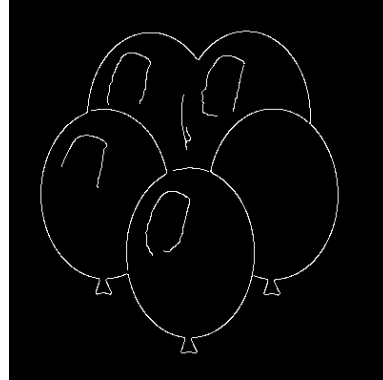
3. Convoluir a imagem  $I$  com a Gaussiana  $G(x)$  ao longo das linhas, gerando uma imagem  $I_x$ , e com  $G(y)$  ao longo das colunas, gerando uma imagem  $I_y$ ;
4. Convoluir  $I_x$  com  $G'(x)$  para gerar  $I'_x$ , e  $I_y$  com  $G'(y)$  para gerar  $I'_y$ ;
5. Calcular a magnitude do gradiente em cada pixel através da Equação 2.9;

$$M(x, y) = \sqrt{I'_x(x, y)^2 + I'_y(x, y)^2} \quad (2.9)$$

6. Supressão não máxima, tratando-se da busca de máximos locais de  $M(x, y)$  nas direções transversais às bordas, reduzindo num mapa de bordas finas;
7. Limiarização adaptativa (histerese), que consiste em uma limiarização baseada em dois limiares,  $\tau_1$  e  $\tau_2$ , sendo um o valor alto de limiar e o outro o valor baixo. Aplicar a limiarização com esses dois limites permite um processo de complementação das descontinuidades da primeira limiarização aproveitando o resultado da segunda.



(a)



(b)

Figura 2.9: Segmentação de imagem utilizando método Canny - (a) imagem em níveis de cinza; e (b) imagem segmentada.

## 2.5 Geração de imagens do sistema nervoso

Existem diferentes técnicas para captura de imagens dos órgãos do corpo humano, em especial do sistema nervoso. Uma técnica de ampla aplicação é a de geração de imagens por ressonância magnética.

Imagem por ressonância magnética é uma técnica que permite criar imagens detalhadas do corpo humano através da utilização de um forte campo magnético e ondas de rádio. Pelo seu grau de detalhamento é utilizada para examinar estruturas mínimas no corpo humano, como ligamentos, cartilagens e estrutura cerebral.

O corpo humano é formado em grande parte por água ( $H_2O$ ), e as moléculas desse composto contêm prótons de hidrogênio que estão em movimento desordenado constante. Este cenário permite a aplicação do seguinte procedimento de ressonância magnética: os prótons são alinhados quando um forte campo magnético é aplicado (o alinhamento consiste em fazer com que os prótons girem no mesmo sentido, movimento chamado de *spin*); após alinhar os prótons, um pulso de onda de rádio é emitido que faz com que os prótons alternem o seu estado de movimento atual; quando esse pulso é finalizado, os prótons iniciam o processo de retorno ao seu *spin* normal e esse processo de retorno produz um sinal de rádio que é captado pelos sensores do equipamento de ressonância magnética, gerando assim a imagem [6].

A Figura 2.10 exibe um exemplo de utilização da técnica acima.

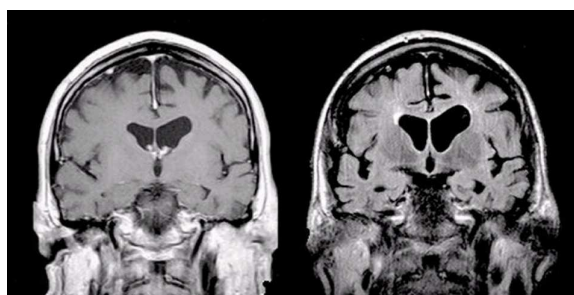


Figura 2.10: Técnica para geração de imagens digitais do corpo humano a partir da utilização de ressonância magnética para avaliação de estágios da doença de Alzheimer<sup>9</sup>.

## 2.6 O arquivo NIFTI

Imagens médicas podem ser geradas em diferentes formatos, cada um com sua especificidade e aplicação [7], e nesse trabalho o tipo de arquivo utilizado foi o tipo *Neuroimaging Informatics Technology Initiative* (NIFTI).

---

<sup>9</sup>Disponível em: Case School of Medicine - <https://bit.ly/2zGdee8>. Acesso: 30/11/2018.

O arquivo NIFTI é um arquivo criado no início dos anos 2000 com o intuito de prover um formato para neuroimagens, em ambiente tri dimensional, que fosse robusto do ponto de vista de representação de pixel (maior quantidade de bits para representação com e sem sinal) e que apresentasse informações úteis para remover ambiguidades no que tange à orientação espacial da imagem [7].

Esse arquivo possui a extensão *.nii* e possui um cabeçalho (espaço com metadados, ou seja, informações de orientação, informações do paciente, etc.) com 352 *bytes*, sendo o tamanho total do arquivo variado, de acordo com o tipo de dado utilizado para representar o volume tri dimensional.

Para realizarmos a visualização da imagem em um plano cartesiano, acessamos o volume NIFTI através de *fatias (slices)*, capturando assim as imagens em um determinado eixo de orientação (sagital, coronal ou axial/transversal, conforme exibido na Figura 2.11) e em uma determinada posição espacial.

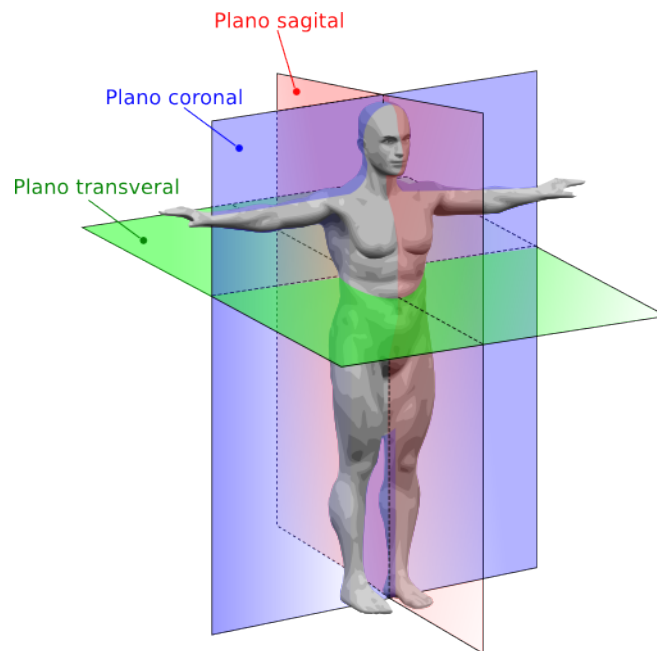


Figura 2.11: Eixos do corpo humano<sup>10</sup>.

## 2.7 Inteligência Artificial

A inteligência artificial (*Artificial Intelligence* (AI)) compreende métodos, ferramentas e sistemas para resolução de problemas que normalmente requerem a inteligência de humanos [8] [9] [10]. Inteligência é um termo definido como a habilidade de aprender de

---

<sup>10</sup>Disponível em: Wikipedia - <https://bit.ly/2Ug1MOH>. Acesso: 30/11/2018.

forma efetiva, reagir de forma adaptada, tomar decisões próprias, comunicar em língua ou imagem de forma sofisticada e de entender os fenômenos e ações do ambiente.

Os direcionamentos para o desenvolvimento de AI são dois: (1) orientar o desenvolvimento de métodos e sistemas que solucionem problemas de forma diferente das soluções realizadas pelos humanos, porém alcançando resultados semelhantes; (2) desenvolver métodos e sistemas modelando a forma com que o humano resolve determinados problemas ou a maneira como o cérebro opera.

O segundo direcionamento, quando da modelagem da ação de um humano em uma situação de resolução de problema, consiste em uma das etapas para a elaboração de um sistema especialista. Um sistema especialista é um programa que pode fornecer expertise (perícia) para resolver problemas em uma determinada área de aplicação da mesma forma que humanos qualificados nessa situação o fazem. Esses sistemas são consistentes na maioria das situações, fáceis de transferir e documentar, e baratos, entretanto são programados sobre um comportamento prévio de um humano, podendo ser frágeis, uma vez que são otimizados para um ambiente específico, incapazes de se adaptar a qualquer mudança do ambiente.

A solução para esse cenário consiste em modelar o funcionamento do cérebro humano como um sistema, de forma que os diferentes tipos de problemas sejam assimilados e processados através de estruturas e variáveis que se adaptam ao longo do tempo. Sistemas desse porte são conhecidos como sistemas com aprendizagem de máquina (*machine learning*).

### 2.7.1 Aprendizagem de Máquina

A aprendizagem de máquina é um método presente em alguns sistemas computacionais, permitindo que tais sistemas evoluam de acordo com os dados processados, reconheçam padrões (classificação de informação em diferentes classes) e tomem decisões com a menor intervenção humana possível.

Para elaborarmos um sistema que aplica aprendizagem de máquina precisamos de três fatores [11]: representação, avaliação e otimização. Precisamos de representação para traduzirmos o mecanismo de reconhecimento (classificação) em uma linguagem formal para o computador através de um algoritmo; avaliação para mensurar a performance da atividade de classificação através de uma função objetiva ou função de pontuação; e por fim, otimização para elevar os resultados obtidos nas funções objetivas.

A representação consiste na etapa de codificação de um sistema capaz de realizar procedimentos de classificação que utilizem técnicas de aprendizagem. A avaliação, também conhecida como etapa de treinamento, consiste em, após elaborar o sistema, executá-lo com uma quantidade relevante de dados para ajustar os seus parâmetros de execução,

através das características dos dados de entrada, de forma que as suas próximas decisões sejam realizadas com menor influência do usuário. Nesse ponto a quantidade de dados a ser utilizada precisa ser relevante, porém ponderada, pois uma grande quantidade de dados pode implicar na ocorrência do efeito de *overfitting*, que ocorre quando o mecanismo de classificação se ajusta somente aos dados do programador, ou seja, só apresentará resultados satisfatórios se executado com os dados utilizados em sua construção/treinamento. A otimização ocorre após a etapa de teste, que consiste em avaliar a performance das decisões nas etapas de classificação [12], com novos dados de entrada, corrigindo eventuais falhas através de um novo conjunto de dados de treinamento ou ajuste de algoritmo.

A etapa de treinamento pode ser realizada seguindo um dos dois paradigmas de aprendizagem computacional: supervisionada e não supervisionada. Na aprendizagem supervisionada o conjunto de dados está rotulado, ou seja, sabe-se previamente a classificação correta para o dado de entrada, fazendo com que o sistema aprenda de forma mais rápida e direta os tipos de classe que ele deve entender e classificar. Para a aprendizagem não supervisionada esses rótulos não são conhecidos, cabendo ao sistema realizar, com maior performance possível, a segmentação e classificação dos dados a partir dos atributos neles presentes.

O objetivo principal da aprendizagem de máquina é generalizar, que é a habilidade do sistema se adaptar e classificar de forma correta os dados desconhecidos. Essa capacidade de generalização pode ser avaliada através de diferentes métricas, sendo algumas delas a acurácia ( $\chi$ ), matriz de confusão, sensibilidade ( $\phi$ ), especificidade ( $\omega$ ), precisão ( $\psi$ ) e *F1 Score* ( $\varrho$ ).

A função de métrica mais simples para a tarefa de classificação é a acurácia de classificação. Ela calcula a fração de amostras que são classificadas corretamente, resultando em um valor entre 0 e 1 (0 indicando classificação errônea para todas amostras e 1 indicando classificação correta para todas amostras). Sua fórmula é dada pela Equação 2.10:

$$\chi = \frac{1}{N} \sum_{i=1}^N 1[y_i == \hat{y}_i], \quad (2.10)$$

onde  $y_i$  e  $\hat{y}_i$  são os rótulos existentes e os obtidos, respectivamente, pela rede em um determinado conjunto de amostras.

A matriz de confusão é uma métrica poderosa para avaliar corretamente modelos de classificação. Para um problema de classificação com  $C$  classes, a matriz de confusão é uma matriz  $C \times C$ , onde o elemento  $M_{ij}$  nessa matriz exibe o número de amostras da classe  $i$  que foram classificadas como classe  $j$  utilizando a rede convolucional. Obviamente,  $M_{ii}$  exibe o número de amostras que foram corretamente classificadas.

Para uma classificação binária, existem 4 resultados: verdadeiro positivo (VP), ou seja,

amostras de uma classe A que foram classificadas como classe A; falso negativo (FN), amostras de uma classe A que foram classificadas como classe B; falso positivo (FP), amostras de uma classe B que foram classificadas como classe A; e verdadeiro negativo (VN), amostras de uma classe B que foram classificadas como classe B. A acurácia para uma matriz de confusão é obtida através da Equação 2.11.

$$\chi = \frac{VP + VN}{VP + VN + FN + FP} \quad (2.11)$$

Para uma classificação de múltiplas classes, temos que para qualquer classe  $i$  na matriz:

$$FN_i = \sum_{j \neq i} M_{ij}, \quad (2.12)$$

fornece o número de amostras falso negativas, e

$$FP_i = \sum_{j \neq i} M_{ji}, \quad (2.13)$$

fornece o número de amostras falso positivas. A acurácia para esse cenário é obtida através da equação Equação 2.14.

$$\chi = \frac{\sum_i M_{ii}}{\sum_i \sum_j M_{ji}} \quad (2.14)$$

A sensibilidade e a especificidade são geralmente utilizadas em classificações binárias, podendo ser utilizadas em classificações de múltiplas classes. A sensibilidade mede a proporção de amostras de uma classe A que foram classificadas como classe A. A especificidade por outro lado mede a proporção de amostras de uma classe B, contrária a classe A, que foram classificadas como classe B. Nos cenários de classificação binária é comum adotarmos uma classe sendo a positiva e a outra como negativa. Suas representações são exibidas nas Equações 2.15 a 2.16.

$$\phi = \frac{VP}{VP + FN} \quad (2.15)$$

$$\omega = \frac{VN}{VN + FP} \quad (2.16)$$

A precisão, também em classificações binárias, mede de forma geral a quantidade de amostras que foram classificadas como uma determinada classe (positiva). Sua formulação é exibida na Equação 2.17.

$$\psi = \frac{VP}{VP + FP} \quad (2.17)$$

Por fim, temos a métrica *F1 Score*, ou pontuação F1, que utiliza a sensibilidade e a precisão para fornecer um valor único e que seja consistente o suficiente para informar um resultado. Dessa forma, a *F1 Score* computa a média harmônica entre essas duas métricas. Para uma classificação de múltiplas classes a *F1 Score* também pode ser utilizada através da média ponderada de uma classe específica. Sua representação para classificações binárias é exibida na Equação 2.18.

$$\varrho = \frac{2 \times VP}{(2 \times VP) + FP + FN} \quad (2.18)$$

Em sistemas de aprendizagem de máquina os dados nunca serão o suficiente pra realizar uma boa generalização, mesmo que já haja vários. O sistema precisa embutir algum conhecimento (viés - *bias*) ou assumir alguns fatores para generalizar o conjunto de dados como um todo de forma mais simples.

Um sistema com aprendizagem de máquina precisa muitas vezes combinar diferentes modelos, por exemplo, redes neurais e *Support Vector Machine* (SVM), ou seja, ao invés de assumirmos um algoritmo de aprendizagem único combina-se alguns algoritmos em busca da melhor resolução de problema possível. Obviamente que o esforço por parte do programador é um pouco maior, mas o benefício é considerável.

## 2.7.2 Redes Neurais Convolucionais

O cérebro é um computador altamente complexo, não linear e paralelo. Ele possui a capacidade de organizar seus constituintes estruturais, conhecidos como neurônios, de forma a realizar certas computações (reconhecimento de padrões, percepção, controle motor, etc.) muitas vezes mais rápidas que um computador digital possa executar hoje.

De forma geral, uma rede neural artificial é um processador distribuído massivamente paralelo feito de unidades simples de processamento, que são naturalmente propensas para armazenar conhecimento experimental e tornar esse conhecimento disponível para uso [13]. Ela se assemelha ao cérebro humano em dois aspectos:

- Conhecimento é adquirido pela rede a partir do seu ambiente através de um processo de aprendizagem;
- As forças das conexões inter neuronais, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

Para entendermos melhor esse tipo de sistema computacional precisamos entender o elemento primário de sua formação, o neurônio. Os neurônios são a base construtora do sistema nervoso (células cerebrais), que se comunicam através de sinapses que são estruturas elementares e unidades funcionais. De forma geral, assume-se que a sinapse é



uma conexão simples que pode impôr excitação ou inibição, mas não ambas no neurônio receptor do sinal.

A plasticidade, termo que refere-se a capacidade do sistema nervoso se adaptar ao ambiente, pode ser alcançada através da criação de novas sinapses entre os neurônios ou na mudança das já existentes. Os neurônios, em suma, são formados por corpo celular, responsável pela produção de proteínas, membranas e respiração celular, axônios, que são linhas de transmissão e os dendritos, as zonas receptivas. A Figura 2.12 exibe uma ilustração dessa estrutura celular.

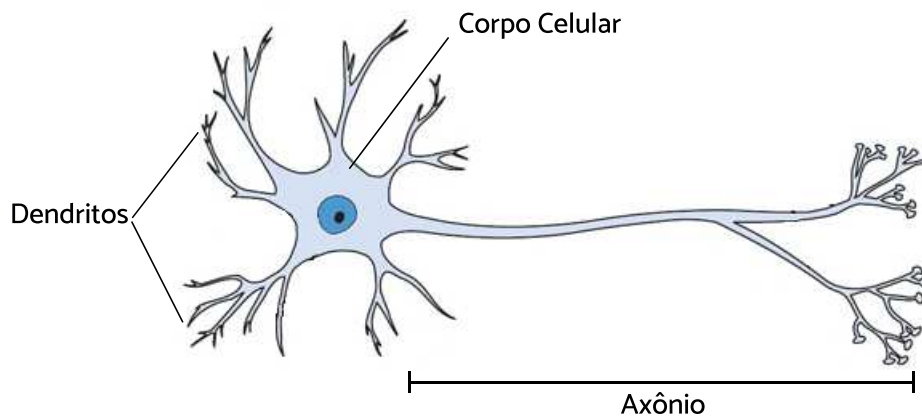


Figura 2.12: Estrutura de um neurônio biológico<sup>11</sup>.

De forma análoga, um neurônio artificial é uma unidade de processamento de informação que é fundamental para a operação de uma rede neural. Um neurônio de uma rede neural artificial é composto, geralmente, por três elementos básicos:

- a) Um conjunto de sinapses, ou conexões, onde cada conexão é caracterizada por um peso ou força. Especificamente, um sinal  $X_j$  na entrada da sinapse  $j$  conectada ao neurônio  $k$  é multiplicada pelo peso sináptico  $w_{kj}$ . O primeiro subscrito refere-se ao neurônio destino e o segundo subscrito refere-se à entrada na sinapse à qual esse peso  $w$  se refere. O peso sináptico de um neurônio artificial pode variar em uma faixa de valores positivos e negativos.
- b) Um somador para somar os sinais de entrada, ponderados pelas respectivas sinapses do neurônio.
- c) Uma função de ativação para limitar a amplitude de saída do neurônio. Tipicamente, a faixa normalizada da amplitude de saída do neurônio é um valor no intervalo  $[0,1]$  ou  $[-1,1]$ .

---

<sup>11</sup>Disponível em: ResearchGate - <https://bit.ly/2Pii5Hb>. Acesso: 30/11/2018.

Na maioria dos modelos também existe um valor de viés externamente incluído, denotado como  $b_k$ . O viés tem o efeito de elevar ou diminuir a entrada de rede da função de ativação, dependendo do seu valor. Em suma, possui o efeito de aplicar uma transformação afim na saída da combinação linear das entradas e pesos sinápticos. Em termos ilustrativos, um neurônio  $k$  possui o formato exibido na Figura 2.13:

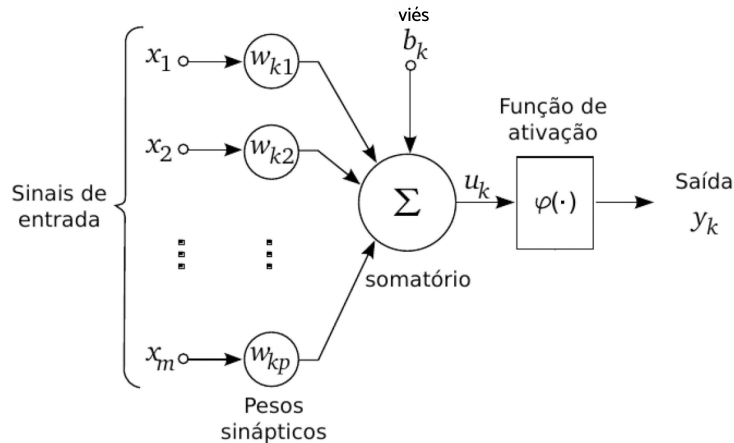


Figura 2.13: Estrutura de um neurônio artificial<sup>12</sup>.

Matematicamente, um neurônio  $k$  é composto pelas Equações 2.19 a 2.20

$$u_k = b_k + \sum_{j=1}^m w_{kj} \times x_j \quad (2.19)$$

$$y_k = \varphi(u_k) \quad (2.20)$$

Apesar da Equação 2.19 iniciar o somatório com  $j = 1$ , considera-se o valor  $x_0$  como sendo igual a 1 e  $w_{k0}$  igual a  $b_k$ .

A Equação 2.20 representa a função de ativação. A função de ativação consiste de uma função que recebe um valor como sinal de entrada e produz uma resposta binária ou um valor em um espectro de valores possíveis. Essa resposta produzida depende do tipo de função, e os tipos mais comuns são: função de limiar, função sigmoide, função tangente hiperbólica, função de ativação linear retificada e função *softmax*.

No funcionamento da função de limiar, a saída do neurônio toma o valor 1 se  $u_k$  for não negativo, caso contrário toma o valor 0. Essa função é exposta na Equação 2.21.

$$\varphi(u_k) = \begin{cases} 1, & \text{se } u_k \geq 0 \\ 0, & \text{se } u_k < 0 \end{cases} \quad (2.21)$$

<sup>12</sup>Disponível em: Scielo - <https://bit.ly/2KV1xHp>. Acesso: 30/11/2018.

Para termos uma função que não haja de forma binária, temos a função sigmoide, que pode fornecer um grau mais livre da real resposta de acordo com os tipos de entrada. A função sigmoide é exibida na Equação 2.22.

$$\varphi(u_k) = \frac{1}{1 + \exp(-au_k)}, \quad (2.22)$$

onde  $a$  é o parâmetro de inclinação, que caso seja variado pode fornecer diferentes curvas.

De forma similar a função sigmoide, temos a função tangente hiperbólica, que realiza o mapeamento da entrada para um valor entre -1 e 1, enquanto que a função sigmoide mapeia a entrada para uma saída entre 0 e 1. Por ter essa diferença nas saídas possíveis essa função é comumente mais utilizada do que a sigmoide, visto que o espectro de saída é maior. A função tangente hiperbólica é exibida na Equação 2.23.

$$\tanh(u_k) = \frac{\sinh(u_k)}{\cosh(u_k)} = \frac{e^{u_k} - e^{-u_k}}{e^{u_k} + e^{-u_k}} \quad (2.23)$$

A função de ativação linear retificada (RELU, do inglês *Rectified Linear Unit*) mapeia um valor de entrada, caso seja positivo, para ele mesmo, caso contrário o valor de saída é zero. Essa função é executada com maior velocidade pelos neurônios, uma vez que somente é executada quando os valores de entrada são positivos. Ela pode ser utilizada de forma combinada para aproximar outras funções. Essa função é exibida na Equação 2.24.

$$f(u_k) = \begin{cases} u_k, & \text{se } u_k \geq 0 \\ 0, & \text{se } u_k < 0 \end{cases} \quad (2.24)$$

Por fim, temos a função *softmax*, que é utilizada em ambientes multi classe e designa um valor probabilístico de uma determinada amostra pertencer à uma determinada classe. A Equação 2.25 exibe a fórmula para essa função.

$$\text{Softmax}(u_k) = \frac{e^{u_k}}{\sum_j e^{u_j}} \quad (2.25)$$

O comportamento da função de limiar, função sigmoide, função tangente hiperbólica e função RELU são exibidas na Figura 2.14.

Agora que conhecemos a estrutura de um neurônio, bem como suas funções de ativação, podemos analisar os tipos de arquitetura de rede. A forma com que os neurônios da rede neural são estruturados está diretamente conectada com o algoritmo de aprendizagem utilizado para treinar a rede. De forma geral, podemos identificar duas classes fundamentais de arquiteturas de rede: redes de camada única com alimentação adiante e redes multicamadas com alimentação adiante.

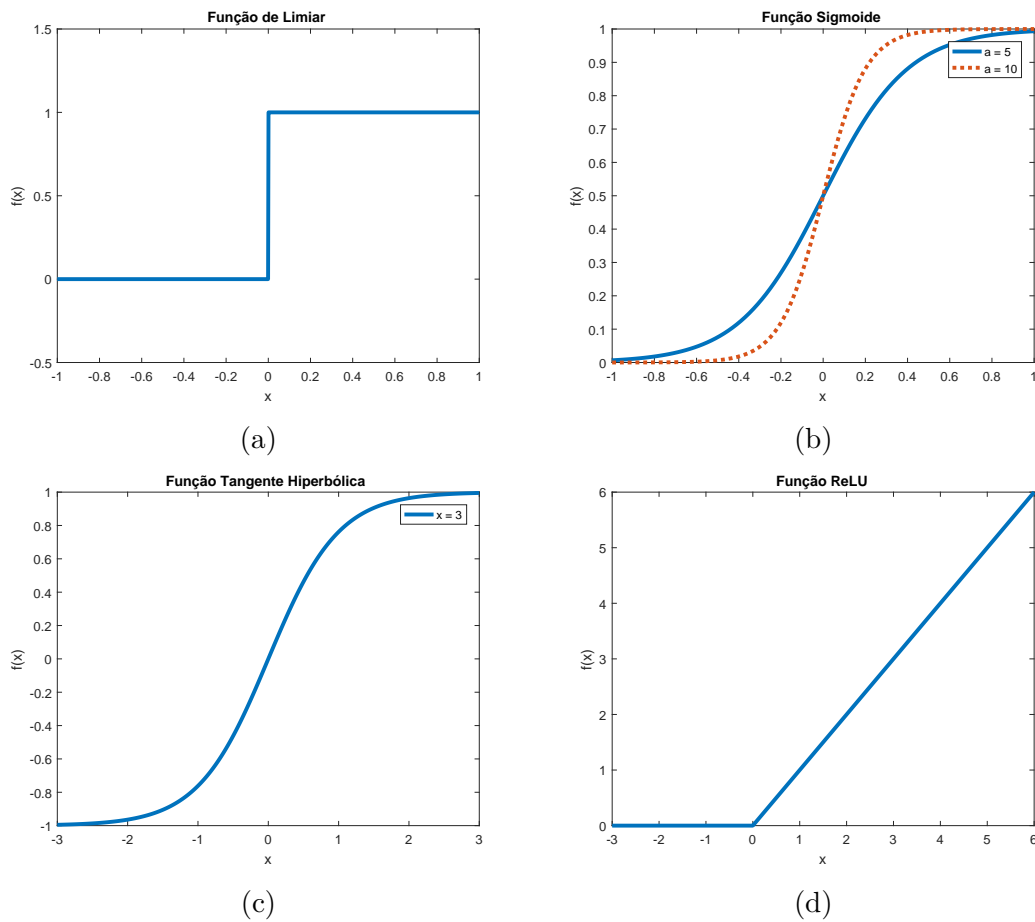


Figura 2.14: Comportamento das funções de ativação - (a) função de limiar; (b) função sigmoide com  $a = 5$  e  $a = 10$ ; (c) função tangente hiperbólica com  $x = 3$ ; e (d) função ReLU.

Na formação mais simples, rede de camada única com alimentação adiante, temos uma camada de entrada de nós de origem que projetam em uma camada de neurônios de saída (nós de computação), mas não vice-versa. Nesse modelo não contamos a camada de entrada de nós de origem porque nenhuma computação é realizada nessa camada. Uma ilustração desse tipo de arquitetura é exibida na Figura 2.15.

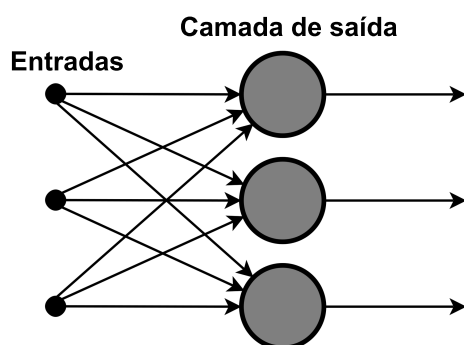


Figura 2.15: Arquitetura de camada única com alimentação adiante<sup>13</sup>.

A segunda arquitetura de rede neural com alimentação adiante, exibida na Figura 2.16, distingue-se pela presença de uma ou mais camadas ocultas, cujos nós de computação são correspondentemente chamados de neurônios ocultos ou unidades ocultas. A função dos neurônios ocultos é intervir entre a entrada externa e a saída da rede de forma útil. Adicionando uma ou mais camadas ocultas, a rede é capaz de extrair características com maior detalhamento.

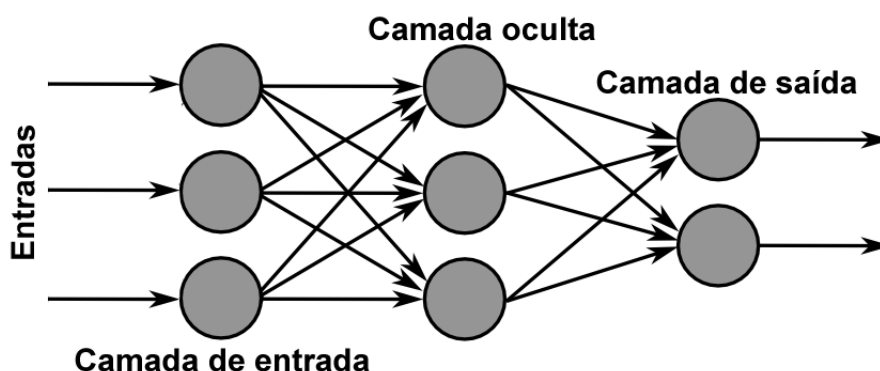


Figura 2.16: Arquitetura de múltiplas camadas com alimentação adiante<sup>14</sup>.

Os nós de origem na camada de entrada da rede fornecem elementos de padrão de ativação (vetor de entrada), que constituem os sinais de entrada aplicados aos neurônios (nós de computação) na segunda camada. Os sinais de saída da segunda camada são utilizados como entradas para a terceira camada, e assim para o resto da rede. O conjunto de sinais de saída dos neurônios da camada final da rede constituem a resposta geral da rede para o padrão de ativação fornecido pelos nós de origem. A rede exibida na Figura 2.16 é dita como totalmente conectada (*fully connected*), uma vez que cada nó em uma camada está conectado com todos os nós da camada seguinte.

<sup>13</sup>Adaptada de: Wikimedia Commons - <https://bit.ly/2rne9vq>. Acesso: 30/11/2018.

<sup>14</sup>Adaptada de: Wikipedia - <https://bit.ly/2KYHzsM>. Acesso: 30/11/2018.

Conhecendo a estrutura de um neurônio artificial, suas funções de ativação e as arquiteturas possíveis precisamos entender o método utilizado para que a rede neural de fato entenda os dados e aprenda, e isso é alcançado através do ajuste dos pesos e viés na rede. No paradigma de aprendizagem supervisionada, um dos algoritmos mais destacados (utilizado neste trabalho) é o algoritmo *backpropagation*, que possui como objetivo reduzir a margem de erro entre o resultado esperado e o resultado obtido na rede neural. A técnica por trás do *backpropagation*, de forma geral, é verificar o quanto a margem de erro varia (gradiente) de acordo com as mudanças nos pesos. Quando uma execução apresenta aumento na taxa de erro ao longo das iterações essa execução é encerrada e os pesos que obtiveram a menor margem de erro entre as saídas obtidas e as saídas esperadas são utilizados como parâmetros finais.

Com a base teórica de redes neurais artificiais, podemos aprofundar o estudo de uma estrutura de rede neural com foco em reconhecimento e classificação de imagens: redes neurais convolucionais.

As redes neurais convolucionais surgem devido ao fato de que ao se utilizar o tipo convencional de redes neurais em imagens a quantidade de neurônios será muito grande, mesmo para arquiteturas de baixa profundidade, tornando impraticável esse tipo de aplicação devido ao custo computacional elevado. A ideia básica das redes neurais convolucionais (*Convolutional Networks* (ConvNets)) é prover uma solução para redução do número de parâmetros permitindo que a rede seja mais profunda com muito menos parâmetros [14].

Em uma rede totalmente conectada, todos os neurônios em uma camada estão conectados com cada neurônio na camada anterior. No caso de imagens em nível de cinza com 8 bits por pixels, a entrada para a primeira camada oculta é uma matriz de  $W$  linhas e  $H$  colunas, com cada pixel da matriz sendo um valor no intervalo  $[0, 255]$ , indicando a intensidade do pixel. Assumindo que existem  $K$  neurônios na primeira camada oculta, cada neurônio nesta primeira camada oculta estará conectado à todos os pixels na imagem em níveis de cinza, conduzindo a  $W \times H$  conexões somente para um neurônio.

Para evitarmos isso realizamos a convolução da imagem de entrada com filtros (regiões quadradas, por exemplo, 3x3, 4x4, etc) de tamanho determinado, reduzindo assim a quantidade de parâmetros. Convoluir uma imagem com um filtro (operação esta que é representada matematicamente pela expressão  $Imagem * Filtro$ ) consiste basicamente em multiplicar uma região da imagem pelo filtro, ponto a ponto, e realizar a soma de todos os produtos, sendo o resultado a convolução para essa região específica, que ainda pode ser aplicado à uma função específica para gerar um valor adaptado (como as funções de ativação descritas anteriormente). De forma geral, se o tamanho da imagem é  $W \times H$  e a camada de convolução é composta por  $L$  filtros de tamanho  $M \times N$ , a saída da camada

de convolução será de  $L$  imagens de tamanho  $W - M + 1 \times H - N + 1$ . A Figura 2.17 ilustra o procedimento de convolução de uma imagem de tamanho 7x7 (área de cor azul claro) com um filtro de tamanho 3x3 (área de cor azul escuro) que gera a imagem de saída de tamanho  $5 \times 5$  (área de cor verde).

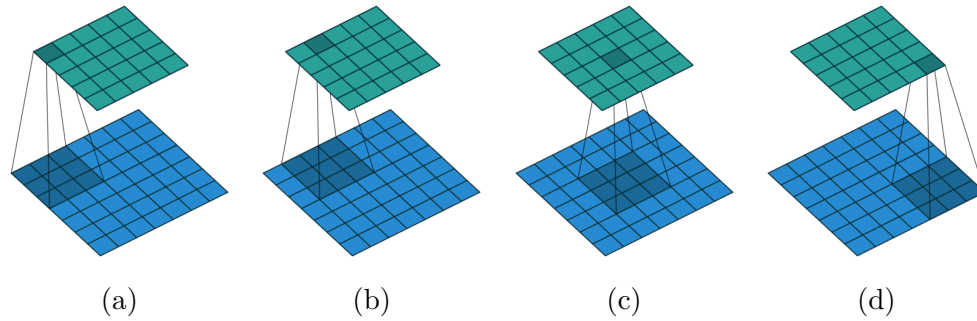


Figura 2.17: Ilustração da operação de convolução - (a) primeira etapa da convolução da imagem com o filtro; (b) segunda etapa da convolução da imagem com o filtro; (c) décima terceira etapa da convolução da imagem com o filtro; e (d) última etapa da convolução da imagem com o filtro<sup>15</sup>.

Uma rede neural convolucional profunda é normalmente composta por mais de uma camada de convolução. Do ponto de vista de processamento de imagens, para imagens em níveis de cinza, um filtro convolucional é um vetor bidimensional (matriz). No caso de imagens multi canais, como imagens no sistema de cor RGB, o filtro de convolução pode ser um vetor bidimensional que é aplicado separadamente em cada canal.

Entretanto, a ideia principal por trás dos filtros convolucionais nas ConvNets é que o resultado da convolução do filtro com entradas multi canais é sempre um canal único. Em outras palavras, se um filtro convolucional  $f$  é aplicado em uma imagem RGB  $X$  com três canais,  $X * f$  precisa gerar uma imagem de canal único (esse canal é único pois as convoluções das três camadas são somadas produzindo uma só). As saídas das camadas de convolução são chamadas de mapas de atributos onde um mapa de atributos é o resultado da convolução de um filtro com a entrada.

Muitas entradas seguem presentes mesmo após a utilização do procedimento de convolução, tornando o processamento da rede ainda custoso. Para isso, utiliza-se um segundo elemento estruturante nas redes neurais convolucionais: a camada de agregação (*pooling*).

A camada de agregação tem como objetivo diminuir progressivamente o tamanho espacial do resultado das convoluções, reduzindo assim a quantidade de parâmetros e a computação exercida pela rede, cooperando também para a diminuição de *overfitting*. No geral, essa camada faz uma subamostragem dos resultados das convoluções, utilizando uma região de tamanho quadrado (por exemplo, 2x2, 4x4, etc.) e um parâmetro  $s$

<sup>15</sup>Adaptada de: Theano - <https://bit.ly/2kkMis3>. Acesso: 30/11/2018.

(deslocamento - *stride*) indicando a quantidade de pixels que a região se movimenta. A Figura 2.18 exibe um exemplo da operação de agregação, onde utilizamos uma janela de tamanho 2x2 e deslocamento de 2 pixels (ao chegar no limite do lado direito da imagem a janela volta ao início, lado esquerdo, e desloca-se para baixo, por 2 pixels), com a função *max()* que captura o maior valor contido na janela em questão.

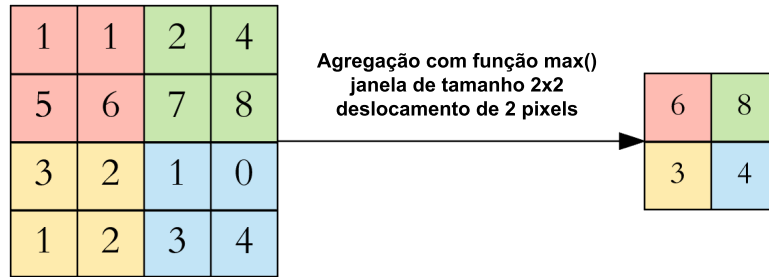


Figura 2.18: Operação de agregação<sup>16</sup>.

Com a redução na quantidade de entradas, podemos agora proceder com a utilização real da rede neural. Essa transição, entre camadas de convolução/agregação com os neurônios da rede neural, é conhecida por camada totalmente conectada, onde cada valor contido no mapa de atributo é fornecido como entrada para a rede neural. A Figura 2.19 exibe um exemplo da arquitetura geral de redes neurais convolucionais.

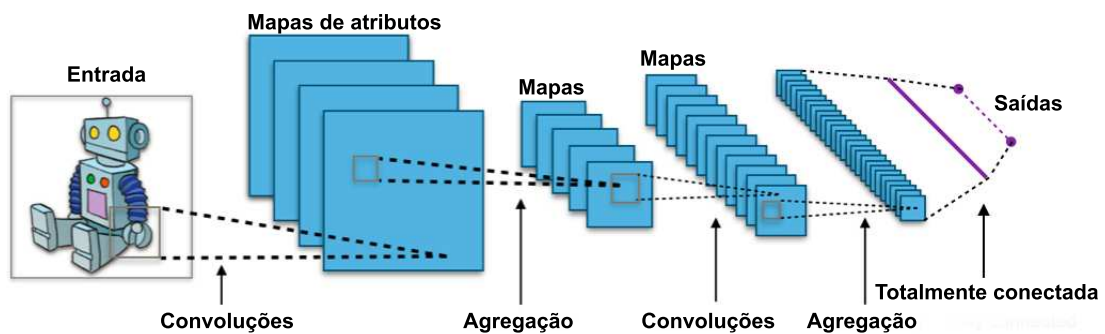


Figura 2.19: Arquitetura geral de uma rede neural convolucional<sup>17</sup>.

## 2.8 Transferência de Aprendizagem

Para que possamos criar uma rede neural convolucional robusta precisamos de dois insumos essenciais: (1) grande quantidade de dados e (2) poder computacional. A demanda pelos dados se dá no sentido de que a rede necessita ser treinada com inúmeras amostras

<sup>16</sup>Adaptada de: Towards Data Science - <https://bit.ly/2QCu72R>. Acesso: 30/11/2018.

<sup>17</sup>Adaptada de: MC.AI - <https://bit.ly/2zIwK9L>. Acesso: 30/11/2018.



de uma determinada informação à ponto de classificar novas entradas de forma robusta e precisa; o poder computacional se deve ao fato de que para processar inúmeras amostras (principalmente quando as entradas são imagens) necessita-se de grande disponibilidade de recursos (memória RAM, processadores multi núcleos, GPUs, etc.) para que o tempo empenhado no treinamento da rede não seja longo. Em grande parte das situações o problema de alta demanda por dados pode ser contornado através da aquisição de novos dados ou pela técnica conhecida como aumento de dados (*data augmentation*), que consiste em gerar novos dados a partir dos já existentes utilizando transformações como rotação, deslocamento e redimensionamento. Entretanto, o aumento da capacidade computacional depende do interesse e possibilidade financeira por parte do projeto, sendo um gargalo estratégico para o desenvolvimento de sistemas com aprendizagem de máquina.

Sendo assim, precisamos de uma abordagem para solucionar a necessidade de recursos computacionais, tornando o tempo de treinamento da rede curto o suficiente para validar o desenvolvimento. Essa abordagem é conhecida como transferência de aprendizagem [1].

Os humanos utilizam transferência de aprendizagem em inúmeras atividades ao longo do dia a dia; reconhecemos uma situação e aplicamos o conhecimento obtido através de experiências anteriores para solucionar essa situação. Quanto maior for a relação entre a situação nova e a experiência obtida, mais fácil será a resolução dessa situação.

Algoritmos de aprendizagem de máquina, tradicionalmente, são elaborados para resolver problemas específicos, o que pode implicar em uma solução de baixa performance. Sendo assim, a transferência de aprendizagem tenta mudar esse paradigma através do desenvolvimento de métodos que transfiram conhecimento de uma ou mais fontes para melhorar o desempenho e/ou reduzir o tempo de treinamento de uma tarefa em específico, conforme exibido na Figura 2.20.

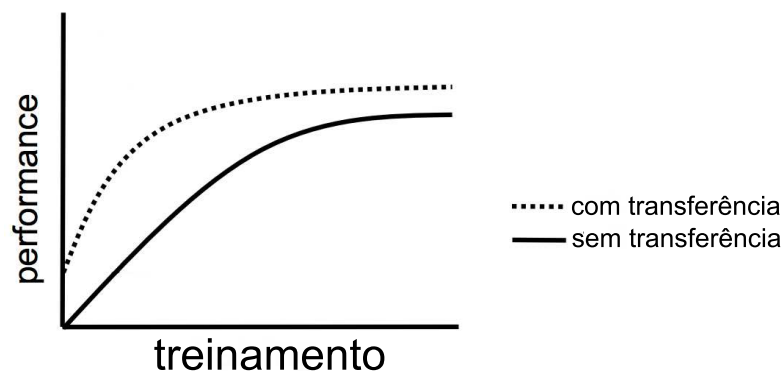


Figura 2.20: Impacto na performance de treinamento com e sem utilização de transferência de aprendizagem. Imagem adaptada de [1].

Para redes neurais convolucionais, existem três cenários principais onde utilizamos a transferência de aprendizagem:

- ConvNets como um extrator de características: nesse cenário utilizamos a estrutura e os pesos prévios da rede neural convolucional, trocando apenas as últimas camadas, pelas camadas de nosso interesse, que são responsáveis pela classificação propriamente dita;
- Ajuste de pesos de ConvNets: nessa situação, além de trocarmos as camadas de classificação de um modelo prévio, realizamos também o retreinamento dos pesos dos neurônios e filtros de convolução, uma vez que queremos tornar o modelo mais adaptado às características dos dados em questão. Esse retreinamento pode ser feito em toda a rede ou em algumas partes, visando manter características que possam ser necessárias (extração de bordas, detecção de regiões, etc.);
- ConvNets pré treinadas: a grande maioria das redes neurais convolucionais levam dias ou semanas para serem treinadas, utilizando grandes recursos computacionais, e isso impede o desenvolvimento ágil de novos classificadores. Sendo assim, o compartilhamento dos pesos da rede é uma prática comum na comunidade de desenvolvedores, auxiliando no alcance de resultados mais rápidos entre os diferentes interesses.

A escolha criteriosa de um cenário é fundamental para que o alcance dos resultados seja prático e robusto, cabendo ao desenvolvedor analisar as características de seu conjunto de dados e optar pela aplicação que efetivamente implique em uma performance considerável.

## 2.9 Trabalhos Correlatos

Em [15] o objetivo está em reconhecer sintomas precoces (pré detecção, pacientes com distúrbios cognitivos leves - *Mild Cognitive Impairment* (MCI)) da doença de Alzheimer, através de imagens de ressonância magnética (*Magnetic Resonance Imaging* (MRI)), para possibilitar a administração correta de drogas que não causem danos irreversíveis ao cérebro. O trabalho consiste de dois experimentos: primeiro, utilizar um método de detecção de Alzheimer, já validado, no estágio de pré detecção e, segundo, propor um novo método com melhor performance, classificando pacientes nos estágios saudável, distúrbios cognitivos leves e Alzheimer. Foram utilizadas imagens de ressonância magnética no eixo coronal, pois exibem com maior facilidade os danos no hipocampo, córtex e ventrículo, de forma pré processada (com melhorias em contraste, ruído e detecção de bordas). No primeiro experimento foi utilizada uma SVM, um modelo de aprendizagem supervisionada que realiza classificação linear binária não probabilística, para classificar as imagens de entrada como pacientes de Alzheimer e pacientes saudáveis. Entretanto, como o foco deste trabalho é pré detectar a doença de Alzheimer (pessoas no estágio de distúrbios cognitivos

leves) esse tipo de técnica não se mostrou ideal devido à SVM ser um método mais propício para classificações binárias (duas classes) [16]. Assim, o segundo experimento consistiu na elaboração de uma rede neural convolucional, uma vez que esse método demonstrou grandes resultados quando utilizado em procedimentos de classificação de imagem [17]. Entretanto, o segundo experimento foi subdividido em duas partes: primeiro, gerar diferentes conjuntos de imagens MRI, com e sem processamento, e posteriormente avaliar a performance da rede com as diversas coleções. Os diferentes conjuntos utilizados para avaliar a rede foram: imagens originais, imagens somente com as bordas dos elementos, imagens com utilização de uma região de interesse (*Region of Interest* (ROI)) com e sem destaque de bordas e imagens com ROI limitada com e sem destaque de bordas. Dois conjuntos de imagens de pacientes foram utilizados, um contendo 1615 imagens (conjunto A) e o outro contendo 1743 imagens (conjunto B). Das 1615 imagens do conjunto A, 1292 foram utilizadas, de forma aleatória, como base de dados para treinamento da rede e 323 para testar o modelo; já para o segundo conjunto o autor não informou a divisão realizada. As melhores performances da rede neural convolucional foram alcançadas utilizando o conjunto de imagens com utilização de uma ROI e sem destaque de bordas, obtendo sensibilidade (taxa de verdadeiros-positivos), de 96% para o conjunto A e de 95% para o conjunto B e especificidade (taxa de verdadeiros-negativos) de 98% para o conjunto A e 98% para o conjunto B.

O trabalho de [18] tem como objetivo propor um método de classificação baseado na combinação de redes neurais convolucionais 3D multi-modelo para capturar várias características das imagens cerebrais originadas por ressonância magnética. O método é composto por duas etapas principais: extração de atributos das imagens 3D por ressonância magnética com as redes convolucionais e classificação das imagens de acordo com a combinação de múltiplas características para diagnóstico de doença de Alzheimer. A vantagem em se utilizar uma combinação de redes convolucionais multi-modelo é: esses tipos de arquitetura conhecidos como aprendizagem profunda (*deep learning*) podem extrair atributos de baixo nível e alto nível, sendo invariante a deslocamentos, mudança de escala e rotações; combinar diferentes modelos amplifica a possibilidade de captura de diferentes características que elevam a performance da classificação final. Duas redes neurais convolucionais foram utilizadas, sendo que uma delas realiza a extração de características através de operações de convoluções entre as imagens do cérebro, em três eixos diferentes, e os filtros convolucionais, havendo o ajuste dos filtros através do algoritmo de *backpropagation*. A outra rede realiza a extração de características de forma diferente, pois para cada etapa de convolução existe uma reconstrução da imagem convoluída e essa reconstrução é analisada através de uma margem de erro entre a imagem original e a reconstruída. Portanto, a segunda rede visa diminuir o erro de reconstrução ajustando os

filtros de convolução, propagando a maior quantidade de características relevantes para as camadas seguintes. Vale ressaltar que a segunda rede opera com diferentes escalas para a mesma imagem 3D do cérebro. As características extraídas por ambas as redes são concatenadas em um vetor de características final para ser classificado como paciente saudável ou paciente de Alzheimer. Para os experimentos deste trabalho, imagens por ressonância magnética de 428 pessoas foram utilizadas, sendo 199 pacientes de Alzheimer e 229 pacientes saudáveis. Como resultado foi obtida acurácia de 88,31%, com sensibilidade de 91,40% e especificidade de 84,42%.

O estudo proposto por [19] consiste na detecção da doença de Alzheimer através da segmentação da região de massa cinzenta utilizando Modelos de Misturas de Gaussianas (*Gaussian Mixture Models* (GMM)) e extração de vetores de pontos baseados em mínimos quadrados parciais (*Partial Least Squares* (PLS)). Esta prática viabiliza a utilização de poucas amostras, que são classificadas posteriormente por SVM. O método baseia-se em determinar uma ROI através da segmentação da massa cinzenta que é afetada pela doença, extrair as características dessa ROI utilizando regressão parcial de mínimos quadrados e classificá-las utilizando uma SVM linear. Primeiramente, realiza-se um pré processamento removendo-se o crânio da imagem cerebral, reajustando o tamanho do tecido cerebral através de operações morfológicas de erosão e dilatação. Após essa etapa, a região da massa cinzenta é segmentada, uma vez que ela é afetada com a progressão da doença. A segmentação neste método é realizada utilizando Modelos de Mistura de Gaussiana, que é um modelo probabilístico que aloca *pixels* para um rótulo específico. Um método semelhante é o *K-Means*, que utiliza um valor médio de distância para alocar os *pixels* para um determinado rótulo de uma classe. Após realizar a segmentação da região de massa cinzenta os atributos remanescente são extraídos através da estimação baseada em algoritmo de mínimos quadrados parciais. Vetores de ponto são extraídos pelo algoritmo e são utilizados como vetores de atributos. Mínimos quadrados parciais é um método estatístico que é utilizado para modelar relações entre variáveis observadas em termos de variáveis ocultas. Com os vetores de ponto obtidos, utiliza-se SVM para treinar um modelo de classificação entre pacientes de Alzheimer e pacientes sem a doença. Foram utilizadas 30 amostras de diferentes pacientes, sendo 15 portadores da doença e 15 saudáveis. Durante a etapa de treinamento foram utilizados 75% das amostras, enquanto que na fase de teste foram utilizadas as 25% restantes. Como resultados foram obtidos sensibilidade de 73,3% e especificidade de 66,6%. Além desses valores, a acurácia do método proposto é avaliada de acordo com o tipo de segmentação utilizada, uma vez que utilizando a segmentação via GMM a acurácia alcança resultados acima dos 80%, enquanto que utilizando o método *K-Means* a acurácia permanece abaixo dos 80%. Outro fator de resultado que vale a pena ser ressaltado é a quantidade de amostras utilizadas

no treinamento, pois o método destaca a dispensabilidade de grandes volumes de dados; nos experimentos realizados a acurácia aumenta de 75% utilizando 4 amostras para 86% utilizando 15 amostras na etapa de treinamento.

A pesquisa executada por [20] fundamenta-se na utilização de redes neurais convolucionais para classificação de pacientes saudáveis, pacientes com distúrbios cognitivos leves (MCI) e pacientes de Alzheimer. A proposta baseia-se em obter imagens MRI no formato *Neuroimaging Informatics Technology Initiative* (NIFTI), pré processá-las e treinar uma rede neural convolucional que fosse capaz de classificar os pacientes em três classes: saudáveis, distúrbios cognitivos leves e Alzheimer. Um total de 900 imagens MRI foram obtidas, sendo 300 de cada classe a ser classificada, utilizando 70% dessa coleção como dados de treinamento e os 30% restante como dados para teste da rede. O pré processamento das imagens consistiu em reconstrução cortical, segmentação volumétrica e extração de imagens no eixo coronal, para que detalhes desnecessários fossem removidos permitindo que a rede fosse treinada de forma mais robusta. As imagens obtidas por ressonância magnética estão no formato NIFTI que fornecem informações do paciente e dados que estruturam as imagens do cérebro nos diferentes eixos e posição temporal. Após o pré processamento, foram extraídas imagens do eixo coronal no formato *Portable Network Graphics* (PNG) dos arquivos NIFTI, em um vetor unidimensional de imagens, e somente as imagens entre os índices 111 e 130, em um total de 250 índices, foram utilizadas, pois cobrem áreas relevantes (com grande probabilidade de danos quando diagnosticada a doença) para a tarefa de classificação. Após a etapa de pré processamento, as imagens extraídas são utilizadas como dados de entrada para a rede neural convolucional, que é uma versão modificada da rede VGG16 de 16 camadas [21], que foi otimizada, com redução no tempo de treinamento, através da inicialização com os pesos da própria rede VGG16 de 16 camadas. Esse modelo de rede contém bons detectores de características bem como detectores de bordas necessários para o trabalho em questão. Nos experimentos foram validados dois tipos de classificação: uma classificação multi classe e uma classificação binária entre as diferentes combinações das classes. Para a classificação multi classe (três classes) foi alcançada uma acurácia de 91,85% enquanto que na classificação binária as seguintes acurácias foram obtidas: 98,33% no cenário Alzheimer x Saudável, 93,89% no cenário Alzheimer x Distúrbio Cognitivo Leve e 91,67% no cenário Distúrbio Cognitivo Leve x Saudável.

O estudo realizado por [22] propõe a construção de redes neurais convolucionais multi nível, de forma cascadeada, para aprender de forma gradual e combinar as características de diferentes modalidades para classificação de doença de Alzheimer utilizando imagens por ressonância magnética e tomografia por emissão de pósitrons (*Positron emission tomography* (PET)) [23]. São duas etapas principais que compõem o método: extração

de atributos através de redes neurais convolucionais 3D (nos três eixos, coronal, axial e sagital), uma para cada modalidade de imagem (MRI e PET) e uma rede neural convolucional 2D para classificação dos pacientes. Duas redes neurais convolucionais 3D são utilizadas para extrair características das imagens, MRI e PET, de forma que sejam invariantes a deslocamentos, escala e rotação. Os mapas de atributos gerados por cada rede neural convolucional 3D são então combinados em um vetor unidimensional, e esses dois vetores unidimensionais são combinados em um vetor bidimensional, elaborando um mapa de características 2D como entrada para a rede neural convolucional final responsável pela classificação. Utilizar uma arquitetura desse tipo, com redes concatenadas, fornece a capacidade de extrair características de alto e baixo nível e possibilita a extração de características anatômicas e funcionais de acordo com a modalidade da imagem. Para validar o método, foram utilizadas imagens de 193 pacientes, sendo 93 pacientes de Alzheimer e 100 pacientes saudáveis, que sofreram um pré processamento (normalização de intensidade e remoção de crânio para as MRI e normalização de intensidade para as PET). Foram executadas etapas de treinamento, validação e teste, dividindo o conjunto de dados em 10 partes e executando combinações com 1 conjunto para teste, 1 conjunto para validação e os demais conjuntos como dados de treinamento. Como resultados foram obtidos os valores de acurácia para as diferentes redes neurais convolucionais 3D para cada modalidade de imagem e para a rede neural convolucional multi modalidade, ratificando o aumento na performance. Para as redes 3D utilizando imagens MRI foi obtida acurácia de 85,47%, enquanto que para a rede utilizando imagens PET foi alcançada acurácia de 87,13%. Para a rede multi modalidade, utilizando ambos tipos de imagem, foi obtida acurácia de 89,64%.

Em [24] propõe-se uma rede neural convolucional que seja discriminativa de forma a classificar 4 tipos de classes: pacientes saudáveis, distúrbios cognitivos leves, distúrbios cognitivos leves em estágio avançado e pacientes de Alzheimer. O método utilizado consiste em um pipeline de duas etapas, sendo elas o pré processamento e o treinamento da rede. A etapa de pré processamento consiste em adquirir imagens por ressonância magnética em um volume 3D NIFTI e realizar correções que elevem a performance da etapa de classificação. São executadas operações de remoção de crânio e segmentação da massa cinzenta, bem como extração de imagens no eixo axial através de um vetor unidimensional de imagens (com utilização de todos os índices com exceção do último e do primeiro por não conterem informação relevante). As redes neurais convolucionais utilizadas neste método são redes já validadas em procedimentos de reconhecimento de objetos, detecção, segmentação e outras atividades. Portanto, a solução consiste em experimentar dois modelos no estado da arte, sendo elas a GoogLeNet [25] e a ResNet [26]. A GoogLeNet difere das arquiteturas convolucionais normais (com muitas camadas em seu início e de

forma sequencial), uma vez que permite processamento em profundidade e de forma paralela, atentando para a possibilidade de *overfitting* caso não haja uma quantidade razoável de amostras. Já na *Residual Network* (ResNet), a ideia principal consiste em aprender características não somente através de operações de convolução, porém utilizar os dados originais para otimizar o funcionamento e elevar a velocidade de convergência. Foram utilizadas duas ResNet nesse trabalho, uma contendo 18 camadas e outra contendo 152 camadas. Para os procedimentos experimentais foram utilizadas imagens MRI de 149 pessoas, sendo 33 pacientes de Alzheimer, 22 com distúrbios cognitivos leves em estágio avançado, 49 em estágio inicial e 45 saudáveis. Para elevar a quantidade de dados, as amostras foram rotacionadas no eixo horizontal, elevando o tamanho do conjunto de imagens para um total de 38024 para as 4 classes. O conjunto de dados foi dividido em 65% para treinamento da rede, 10% para validação e 25% para teste. Os resultados foram analisados em torno da acurácia da classificação realizada por cada rede. A GoogLeNet alcançou acurácia de 98,88%, enquanto que a ResNet, com 18 camadas, obteve 98,01%, e a com 152 camadas obteve 98,14%.

# Capítulo 3

## Solução Proposta

Tomando como base as informações descritas no capítulo anterior, podemos proceder com a abordagem deste trabalho.

A solução proposta consiste em um algoritmo computacional que realiza a avaliação de imagens MRI e consequentemente classifica-as em um estágio da doença de Alzheimer, sendo eles: saudável, portador de distúrbios cognitivos leves e portador da doença de Alzheimer.

O algoritmo recebe imagens digitais provenientes de ressonância magnética, extrai características dessas imagens (bordas, regiões, intensidade de pixels, etc) e de acordo com o conjunto dessas características classifica as imagens, entre os estágios possíveis previamente citados, através de um valor probabilístico.

A descrição teórica de cada etapa da solução é realizada nas seções subsequentes.

### 3.1 Base de dados

As imagens médicas utilizadas para diagnosticar inúmeros quadros clínicos são obtidas através de equipamentos médicos com um funcionamento específico conforme descrito na Seção 2.5. Essas imagens ficam muitas vezes sob porte do paciente e/ou do médico, sendo inutilizada para procedimentos de pesquisa na maioria das vezes. Com esse cenário surgiu a *Alzheimer's Disease Neuroimaging Initiative* (ADNI) no ano de 2004, fundada como uma parceria público-privada. A ADNI é um centro de estudo norte americano elaborado para desenvolver dados clínicos, imagens, dados genéticos e biomarcadores<sup>1</sup> para a detecção precoce e monitoramento da doença de Alzheimer.

A ADNI passou por diferentes estágios de desenvolvimento desde sua fundação, sendo que ao longo desses anos inúmeros participantes foram selecionados para participar dos

---

<sup>1</sup>uma característica que é objetivamente mensurável e avaliada como um indicador de um processo normal biológico, processo patogênico ou uma resposta farmacológica à uma intervenção terapêutica.



procedimentos de pesquisa. Atualmente, indivíduos que acessam o *site* da ADNI se deparam com a informação de que pessoas com idade entre 55 e 90 anos podem estar aptas a participar da pesquisa, sendo realizada de forma gratuita e sem administração de medicações.

## 3.2 Pré processamento das imagens

As imagens obtidas na base ADNI são arquivos no formato NIFTI, ou seja, um arquivo que contém informações de representação tri dimensional (representado por fatias nos eixos coronal, sagital e axial/transversal) de um determinado órgão ou região a ser obtida digitalmente por um procedimento de, por exemplo, ressonância magnética.

Para execução deste trabalho foram tomadas as imagens no eixo coronal, uma vez que é o eixo que fornece informações mais propícias para a classificação do paciente em um dos estágios da doença, pois as três regiões que sofrem grande avaria (córtex, hipocampo e ventrículos) são visíveis de forma simultânea. Após a escolha do eixo precisamos definir quais os índices das fatias que serão utilizadas como dado válido, uma vez que cada índice representa uma posição espacial do volume tri dimensional, conforme exibido na Figura 3.1. Para isso, analisamos que cada volume tri dimensional pode conter 192 ou 256 fatias no eixo coronal, e dessa forma os índices que apresentarão informações de maior relevância são os índices 96 e 128, justamente as fatias centrais, que contém informações úteis dos ventrículos, hipocampo e córtex.

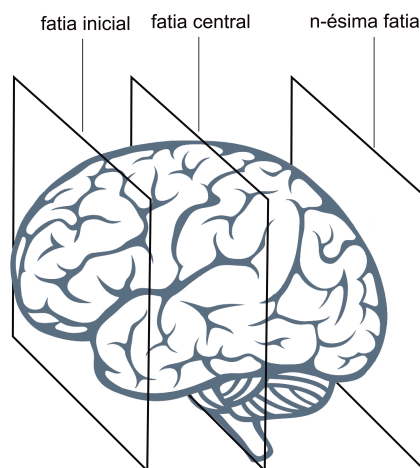


Figura 3.1: Fatias do volume tri dimensional<sup>2</sup>.

---

<sup>2</sup>Adaptada de: KissPNG - <https://bit.ly/2FYkHuv>. Acesso: 30/11/2018.

Entretanto, fatias próximas as fatias centrais também apresentam informação relevante e dessa forma optamos por utilizar as fatias entre os índices 111 à 131 e 83 à 103, para volumes de 256 e 192 fatias respectivamente.

De posse das fatias necessárias, aplicamos procedimentos de filtragem (filtro de mediana) e aprimoramento da qualidade da imagem (*power-law*) para que ruídos e elementos não desejáveis (por exemplo, baixo contraste e excesso de brilho) sejam corrigidos. O fator  $\gamma$  utilizado foi de 0,77, uma vez que as imagens apresentavam demasiado brilho e necessitávamos de um brilho menor para correta aplicação do filtro de mediana.

Após a filtragem da imagem, realizamos a determinação de uma região de interesse (ROI), que visa remover detalhes desnecessários da imagem e elevar a velocidade de processamento por parte da rede neural convolucional. Um exemplo de imagem processada com a região de interesse é exibida na Figura 3.2.

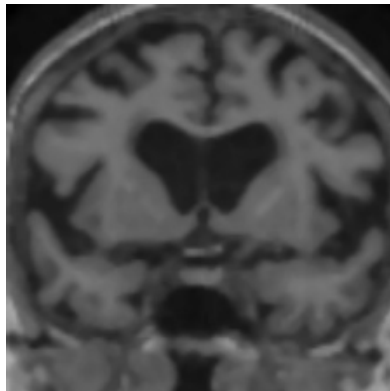


Figura 3.2: Imagem após aplicação de filtro de mediana, transformação *power-law* e determinação de ROI.

Por fim, para critérios de análise de desempenho, um conjunto de dados com segmentação *Canny* foi gerado (exemplo exibido na Figura 3.3), sendo analisada sua performance no Capítulo 4.

### 3.3 Extração de características via transferência de aprendizagem

Com o conjunto de dados em mãos, podemos então, fornecer essas imagens como entrada para a rede neural convolucional. Entretanto, o tempo gasto para treinar uma rede convolucional profunda por completo com uma grande quantidade de imagens seria muito grande. Dessa forma, precisamos de uma solução prática e robusta que execute a captura das características das imagens que serão utilizadas pela rede para propriamente classificar os dados, e essa abordagem é a transferência de aprendizagem.



Figura 3.3: Imagem após aplicação de filtro de mediana, transformação *power-law*, determinação de ROI e segmentação com método *Canny*.

Existem várias redes neurais convolucionais elaboradas em competições (como, por exemplo, o *ImageNet Large-Scale Visual Recognition Challenge* (ILSVRC)) ou trabalhos acadêmicos e que foram compartilhadas entre estudantes e pesquisadores para promover a difusão de redes convolucionais profundas. Dentre as redes existentes, utilizou-se nesse trabalho a rede *Visual Geometry Group - 16 Layers* (VGG16), uma vez que as características para o diagnóstico de Alzheimer são minuciosas, necessitando de uma arquitetura pautada na utilização de filtros de convolução de baixa dimensão ( $3 \times 3$ ), fazendo com que o processamento seja mais minucioso do que quando utilizamos filtros maiores [2][27]. A Figura 3.4 exibe uma evolução das redes vitoriosas na competição da ILSVRC e o crescimento no número de camadas nas redes.

A proposta da utilização da VGG16 neste trabalho é servir como um extrator de características que serão fornecidas para a rede neural. Sendo assim, as camadas totalmente conectadas dessa arquitetura foram removidas e em contrapartida novas camadas foram adicionadas para confeccionar o novo classificador.

### 3.4 Classificação da imagem

Por fim, temos a rede neural, que utiliza os atributos extraídos das imagens e processa-os de forma a identificar padrões e minúcias que descrevem cada classe em questão. A etapa de classificação, realizada de forma supervisionada, consiste em fornecer uma imagem de tamanho  $224 \times 224$  pixels de um paciente e obter um vetor de três valores contendo as

### Evolução da profundidade das redes neurais convolucionais

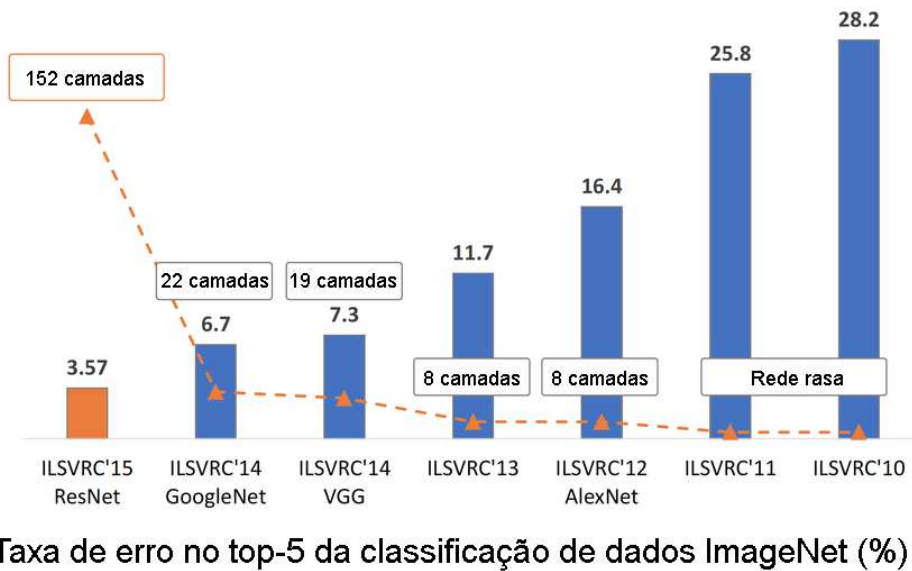


Figura 3.4: Evolução das redes neurais convolucionais profundas. Imagem adaptada de [2].

probabilidades daquela imagem ser de algum estágio (saudável, distúrbio cognitivo leve ou paciente de Alzheimer).

A arquitetura utilizada consiste em uma rede com alimentação adiante (*feed-forward*) de 3 camadas, sendo 1 camada de entrada, 1 camada oculta e 1 camada de saída. A função de ativação na primeira camada é a tangente hiperbólica; na segunda camada a ativação linear retificada e na última camada a função *softmax*, que indica uma probabilidade de o dado de entrada pertencer à alguma das classes.

Por questões de robustez, mecanismos de regularização e prevenção de *overfitting* foram utilizados, sendo eles *Dropout* [28] e Regularização L2 [29]. O mecanismo *Dropout* consiste em, de forma aleatória através de um valor probabilístico, remover algum(uns) neurônio(s) de um ciclo de treinamento. Isto é, esse(s) neurônio(s) não será(ão) utilizado(s) para computar a atualização dos pesos da rede e nem vai(ão) ser atualizado(s) com o algoritmo de *backpropagation* durante este ciclo. Já a regularização L2 consiste em adicionar uma penalidade à função de perda para que a minimização dessa função não ocorra de forma abrupta a ponto de causar *overfitting*.

A função de perda utilizada na rede é a entropia categórica cruzada (*categorical cross-entropy*) ( $\rho$ ) e a função de otimização é a gradiente descendente estocástico (*Stochastic Gradient Descent* (SGD)) [30].

A entropia categórica cruzada mede a performance de um modelo cuja classificação seja

realizada através de um valor probabilístico entre 0 e 1, sendo exibido seu comportamento padrão na Figura 3.5. Para um classificador binário, essa função pode ser calculada através da Equação 3.1, enquanto que para um classificador multi classes essa função pode ser calculada através da Equação 3.2:

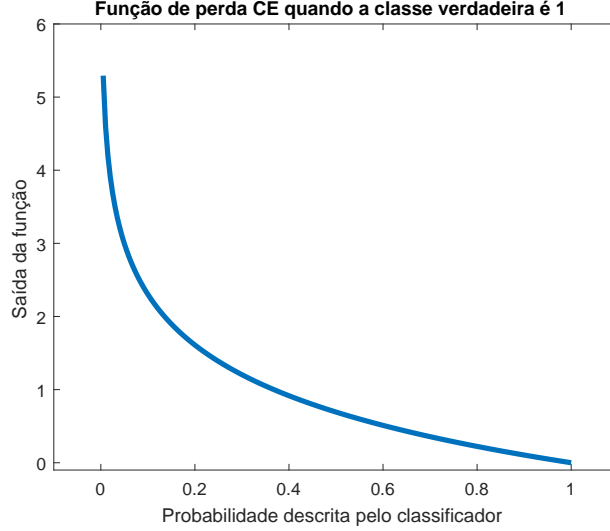


Figura 3.5: Exemplo de comportamento da função de entropia categórica cruzada.

$$\rho_{\text{binário}} = -(y \times \log(p) + (1 - y) \times \log(1 - p)), \quad (3.1)$$

onde  $y$  representa a classe verdadeira e  $p$  a probabilidade descrita pelo classificador de uma determinada entrada ser da classe  $y$ .

$$\rho_{\text{multiclasse}} = -\sum_{c=1}^M y_{o,c} \times \log(p_{o,c}), \quad (3.2)$$

onde  $M$  representa o número de classes,  $y$  um indicador binário (0 ou 1) se a classe  $c$  é a classificação correta para a amostra  $o$ , e  $p$  é a probabilidade descrita pelo classificador da amostra  $o$  ser da classe  $c$ .

O gradiente descendente estocástico é utilizado como função de otimização e sua diferenciação para o gradiente descendente convencional é que nele computamos o gradiente a cada amostra de treinamento, ao contrário do que o gradiente descendente realiza, quando computa o gradiente somente após calcular o erro geral envolvendo todas as amostras de treinamento. Essa função ainda é combinada com uma técnica de aceleração na convergência do gradiente, conhecida como *Nesterov momentum*, podendo ser aprofundada pelo leitor em [31].

Por fim, temos na Figura 3.6 a estrutura geral da rede neural convolucional estabelecida para esse trabalho. A entrada da rede é uma imagem de tamanho  $224 \times 224$  pixels; nas

operações de convolução e agregação, os valores  $y \times y$  indicam o tamanho dos filtros utilizados, enquanto o parâmetro  $S$  na agregação significa o deslocamento, ficando a definição da quantidade de neurônios para o Capítulo 4.

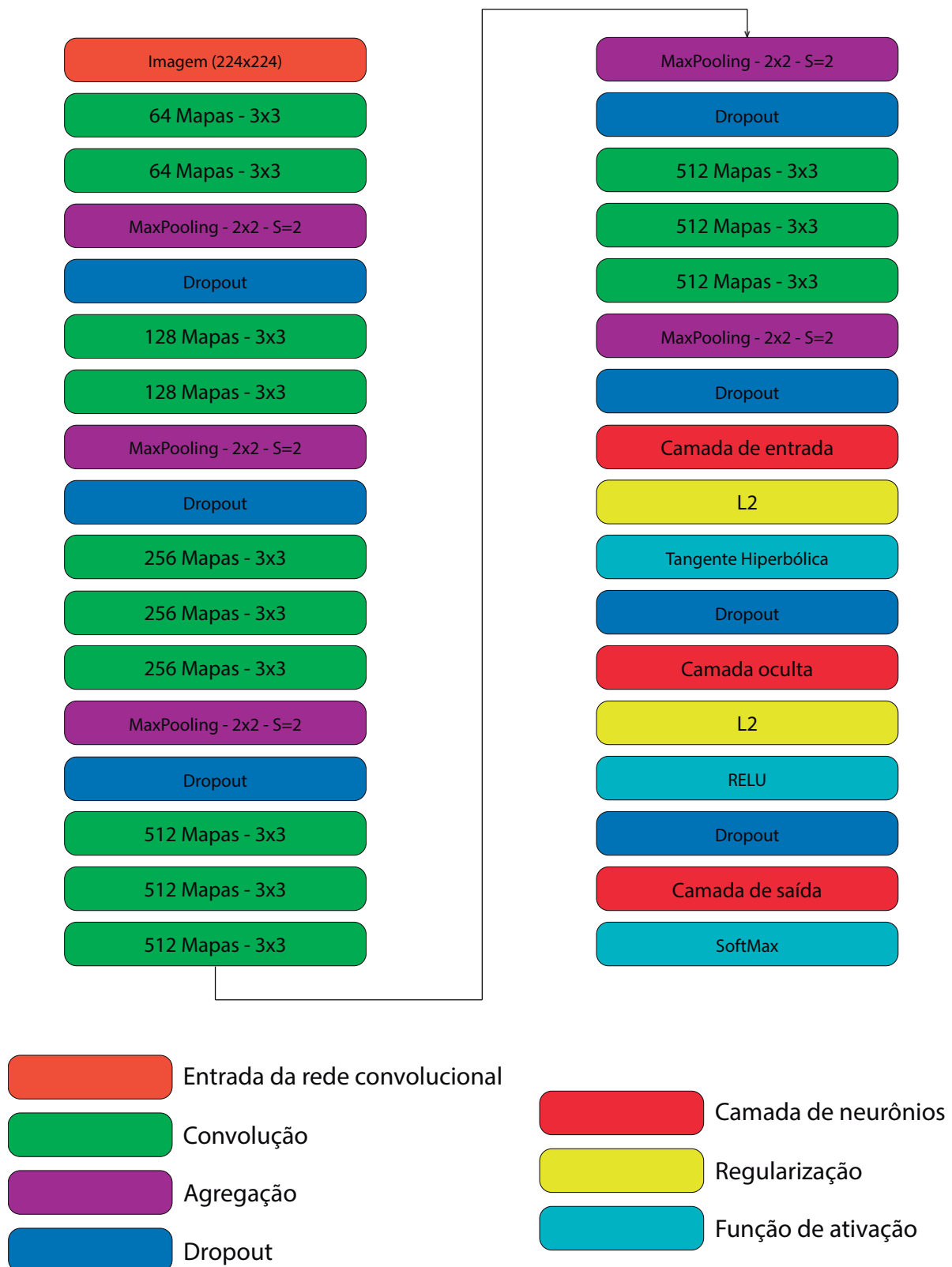


Figura 3.6: Estrutura geral da rede neural convolucional.

# Capítulo 4

## Resultados Experimentais

Visando avaliar a validade do classificador elaborado com a rede neural convolucional, foram executados alguns experimentos e cenários de classificação.

Os dados utilizados nesse trabalho, provenientes da base de dados ADNI, são imagens MRI de indivíduos já diagnosticados em algum estágio da doença. Os estágios da doença, também conhecidos como grupos de pesquisa, são diversos na ADNI, entretanto, para simplificar a análise desse trabalho, iremos utilizar apenas pacientes das seguintes classes: saudável, distúrbios cognitivos leves e portadores da doença de Alzheimer. Cada um desses estágios é descrito da seguinte forma:

- Saudável (S): indivíduos que não apresentam perda de memória, problemas com orientação motora ou problemas em atividades domésticas/profissionais;
- Distúrbios cognitivos leves (D): indivíduos com perda moderada de memória, moderada desorientação temporal e motora e dificuldades nas atividades diárias;
- Portador da doença de Alzheimer (A): indivíduos com perda grave de memória, perda grave de movimentos e perda grave de fala.

Para a condução dos experimentos, foram adquiridas 300 imagens de ressonância magnética, 100 de cada estágio. Dessas 300 imagens, foram obtidas 20 fatias de cada, resultando em um total de 6000 imagens digitais.

A base de dados ADNI utiliza, além das imagens por ressonância magnética, exames clínicos, biomarcadores e dados demográficos para convergir à um diagnóstico mais preciso. Dessa forma, algumas imagens, mesmo sendo de um paciente saudável, podem apresentar semelhanças visuais com uma classe diferente, o que dificulta o trabalho de nosso classificador. Portanto, foram estabelecidos três ambientes de experimentação: (1) conjunto de todas as imagens obtidas na base; (2) conjunto de dados reduzido, com a remoção de determinadas imagens que foram apontadas pelo próprio classificador como



demonstrando muita similaridade com mais de uma classe (uma vez que o resultado do classificador são probabilidades de uma imagem pertencer à uma determinada classe); (3) conjunto de dados reduzidos com aplicação de segmentação *Canny*. A Figura 4.1 exibe um exemplo de cada classe em análise, e nela percebemos, por exemplo, moderada semelhança entre pacientes com distúrbios cognitivos leves e pacientes diagnosticados com Alzheimer.

Além do classificador multi classe, cenários de classificação binária entre as classes foram realizados (Alzheimer x Saudável; Alzheimer x Distúrbios Cognitivos Leves e Distúrbios Cognitivos Leves x Saudável), com intuito de validar o classificador em diferentes modos de operação.

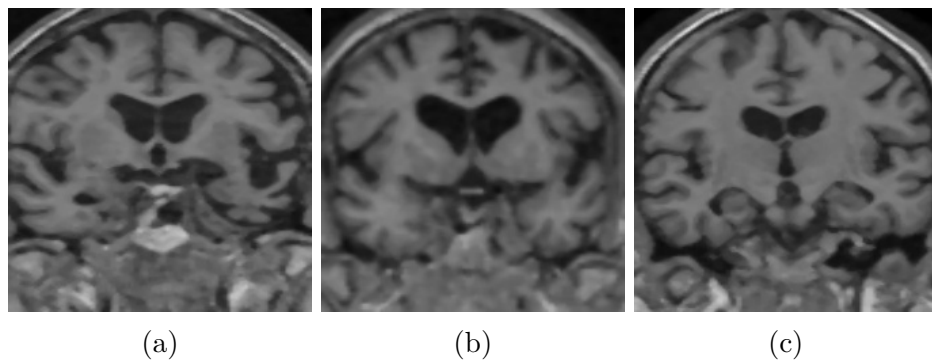


Figura 4.1: Exemplos de imagens utilizadas como entrada para a rede convolucional - (a) portador da doença de Alzheimer; (b) paciente com distúrbios cognitivos leves; e (c) paciente saudável.

## 4.1 Elaboração e treinamento de rede neural convolucional

Para elaborarmos a rede neural convolucional procedemos com a seguinte abordagem:

1. Utilização da rede convolucional VGG16 como extrator de características;
2. Remoção das últimas camadas da rede VGG16 original;
3. Adição de primeira camada (camada 1), contendo 1 neurônio;
4. Adição de segunda camada (camada 2), contendo 1 neurônio;
5. Adição de camada de saída, contendo 3 neurônios;
6. Execução da rede neural convolucional. Ao obter resultados, incrementa-se a quantidade de neurônios na camada 1 e faz com que a quantidade de neurônios na camada

2 seja a quantidade na camada 1 menos uma unidade. Essa etapa é realizada até alcançarmos um cenário de *overfitting*;

7. Após alcançar *overfitting*, utilizar mecanismos de regularização L2 e *Dropout* para ajustar esse cenário. A regularização L2 é inicializada com valor 1 e é dividida por 10 em cada nova execução. O *Dropout* é inicializado em 0.1 e é incrementado em 0.1 em cada nova execução. Ambos procedimentos são realizados de forma simultânea até alcançar a melhor performance.

Por fim a rede neural convolucional consiste de 6 neurônios na camada de entrada, com regularização L2 com penalidade de 0.001 e *Dropout* com valor 0.4; 5 neurônios na camada oculta, com regularização L2 com penalidade de 0.001 e *Dropout* com valor 0.4; e 3 neurônios na camada de saída (sendo alterada para 2 neurônios quando em cenários de classificação binária). O parâmetro *learning rate*, parâmetro utilizado na computação do gradiente descendente estocástico, foi fixado no valor 0.0035, definido empiricamente, pois demonstrou melhor otimização da função de perda.

A rede neural foi executada com 50 iterações (*epochs*) por execução, sendo que a cada iteração os pesos da rede são salvos e somente o melhor conjunto de pesos (que apresentou menor taxa de perda) é utilizado na etapa de testes.

Para o segundo e terceiro ambiente de experimentação foram removidas as imagens que apresentaram menos de 55% de probabilidade de pertencer à uma determinada classe, ou seja, imagens com muitas semelhanças com duas classes diferentes e que impede a melhor performance do classificador. O valor de 55% foi adotado de forma empírica. A diferença entre o segundo e terceiro ambiente é a aplicação da segmentação *Canny* no conjunto de dados a ser utilizado no terceiro ambiente.

## 4.2 Resultados

Após a aquisição do conjunto de dados e elaboração da rede neural convolucional, procedemos com a execução da rede nos diferentes ambientes e cenários de classificação.

Para cada cenário de classificação foram obtidos dois gráficos (acurácia e perda ao longo das iterações para o conjunto de treinamento e validação) e a matriz de confusão para aquele conjunto de dados utilizado. Sobre a matriz de confusão, podemos então calcular os níveis de acurácia, sensibilidade, especificidade, precisão e a pontuação F1.

Nos gráficos que indicam a acurácia ao longo das iterações, o eixo das ordenadas representa a acurácia em porcentagem e o eixo das abcissas representa a iteração corrente. Já nos gráficos que indicam a perda ao longo das iterações, o eixo das ordenadas indica a perda (através da entropia categórica cruzada) e o eixo das abcissas indica a iteração

corrente. Nas matrizes de confusão, as classes são indicadas pelas letras: (A) Alzheimer; (D) Distúrbios cognitivos leves; e (S) Saudáveis. A Tabela 4.1 exibe os resultados experimentais obtidos em todos os ambientes e cenários de classificação, e serão detalhados a seguir.

Resultados						
Ambiente	Classificação	Métricas (%)				
		Acurácia	Sensibilidade	Especificidade	Precisão	F1-Score
Primeiro	Multi Classe	36,00	36,00	68,00	28,60	31,88
	A x S	64,00	98,00	30,00	58,33	73,13
	A x D	54,00	77,00	31,00	52,74	62,60
	D x S	45,00	7,00	83,00	29,17	11,29
Segundo	Multi Classe	36,88	32,78	66,33	23,99	27,70
	A x S	83,33	66,67	100,00	100,00	80,00
	A x D	60,00	100,00	0,00	60,00	75,00
	D x S	48,00	40,00	53,33	36,36	38,10
Terceiro	Multi Classe	65,63	60,83	81,83	63,99	62,37
	A x S	85,00	76,67	93,33	92,00	83,64
	A x D	60,00	100,00	0,00	60,00	75,00
	D x S	86,00	77,50	91,67	86,11	81,58

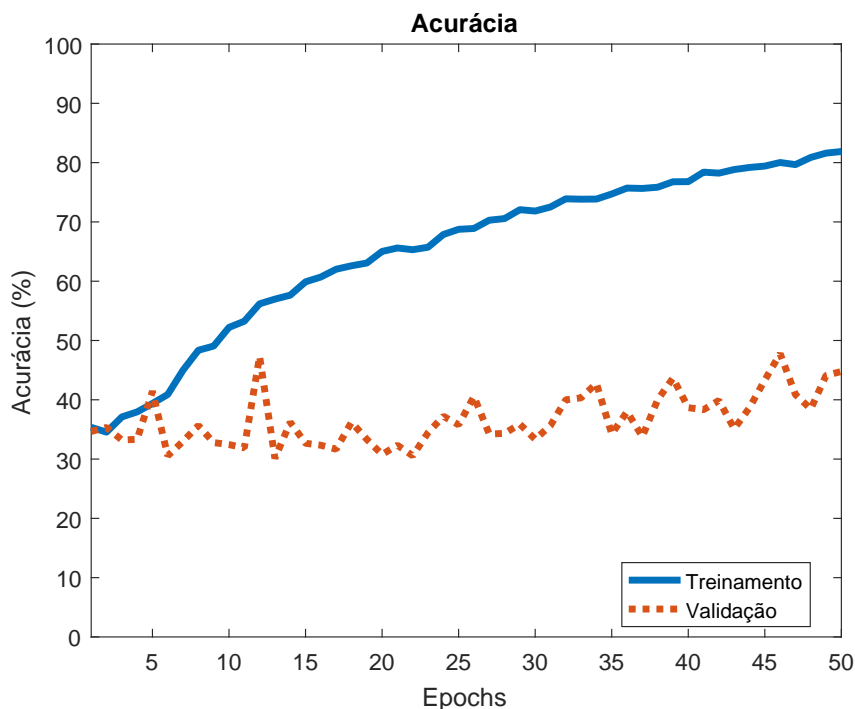
Tabela 4.1: Resultados experimentais.

### 4.2.1 Primeiro ambiente: conjunto com todas imagens

#### Classificação multi classe

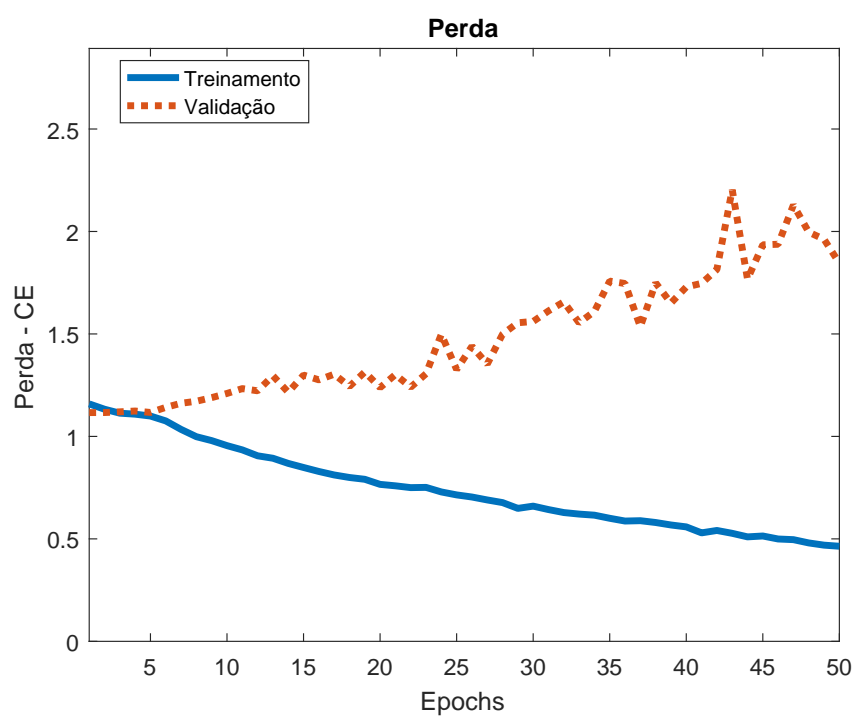
A classificação multi classe no primeiro ambiente de experimentação consiste em realizar a classificação das três classes em análise nesse trabalho, utilizando-se de todas as imagens obtidas para o conjunto de dados. Ao total, foram utilizadas 6000 imagens (2000 de cada classe), divididas na seguinte proporção: 80% para treinamento, 15% para validação e 5% para teste. Os resultados podem ser visualizados na Figura 4.2.

Analisando os gráficos de acurácia e perda, percebe-se a ocorrência de *overfitting*, uma vez que a taxa de perda, para as amostras do conjunto de validação, eleva-se, enquanto que para as amostras do conjunto de treinamento essa taxa diminui.

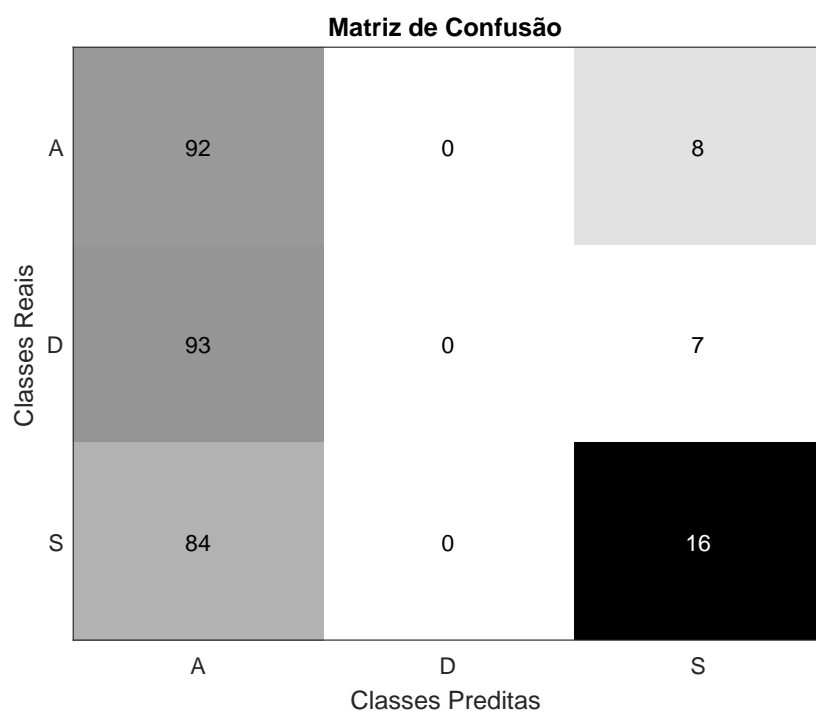


(a)

Figura 4.2: Classificação multi classe no primeiro ambiente de experimentação - (a) Gráfico de acurácia x iterações;



(b)



(c)

Figura 4.2: Classificação multi classe no primeiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

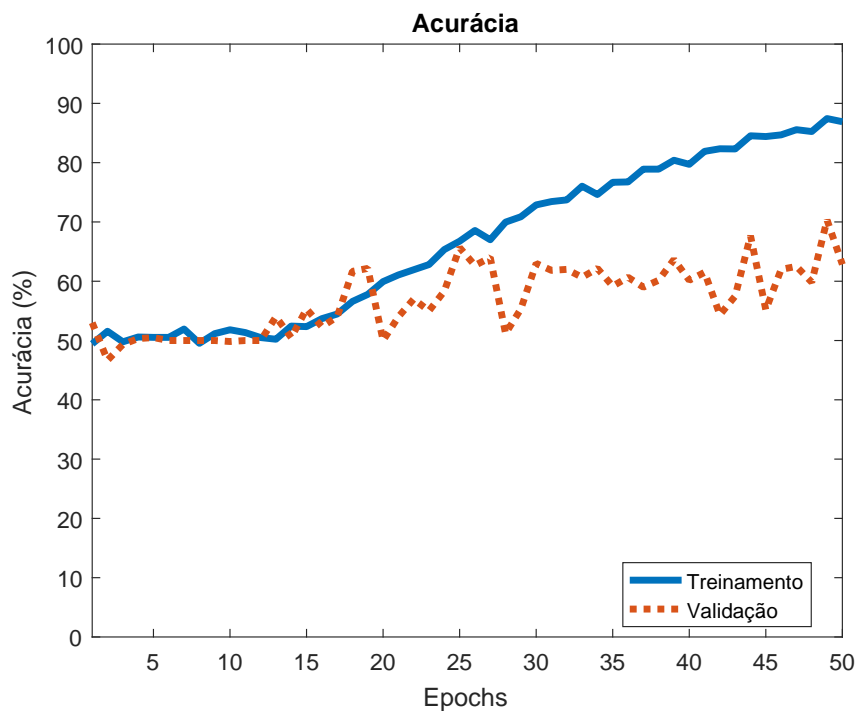
Nesse cenário, percebe-se que o classificador não obteve parâmetros suficientes para diferenciar as três classes de forma robusta, designando a maioria das imagens como

portadores da doença de Alzheimer.

### Classificação binária - Alzheimer x Saudável

A primeira classificação binária no primeiro ambiente consistiu em classificar pacientes portadores da doença e pacientes saudáveis. Ao total, foram utilizadas 4000 imagens (2000 imagens de cada), sendo divididas na mesma proporção do cenário anterior. Os resultados podem ser visualizados na Figura 4.3.

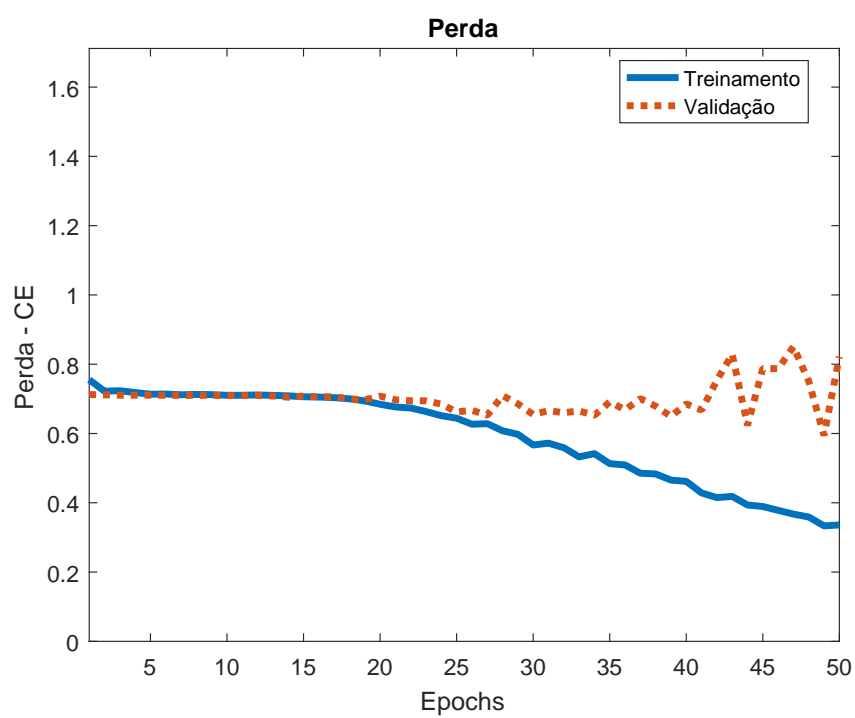
Avaliando os gráficos obtidos nesse cenário, percebe-se que até a iteração 25 as curvas de acurácia e de perda, de ambos conjuntos, apresentaram o mesmo comportamento. A partir dessa iteração a rede já não consegue obter maiores detalhes passíveis de distinção para as classes em análise.



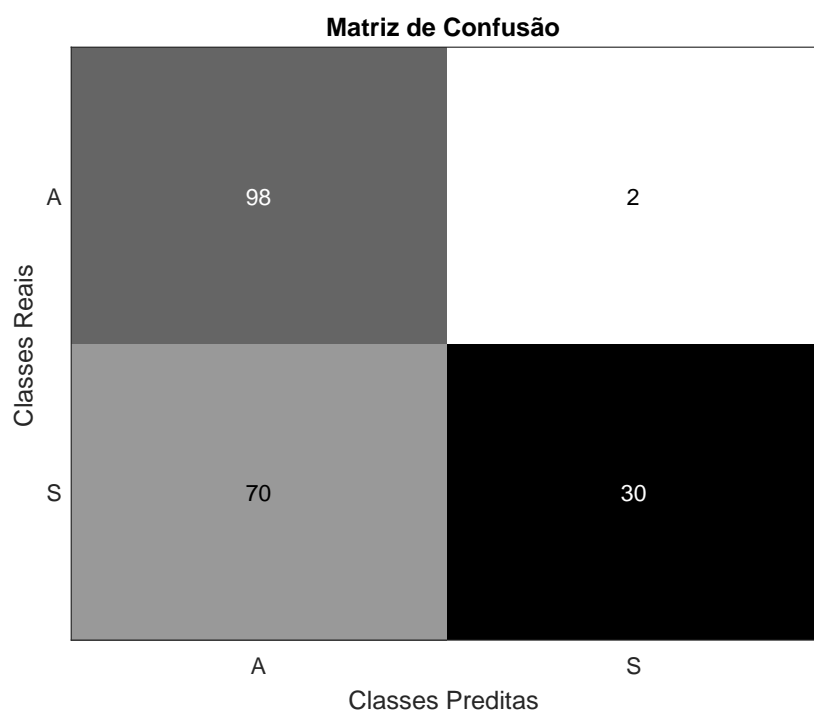
(a)

Figura 4.3: Classificação binária (Alzheimer x Saudáveis) no primeiro ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Percebe-se uma melhora nos resultados deste cenário, uma vez que as imagens dessas classes, em sua maioria, apresentam diferenças consideráveis. Entretanto, a classificação ainda assim converge para uma alta taxa de falsos positivos (pacientes saudáveis classificados como portadores da doença).



(b)



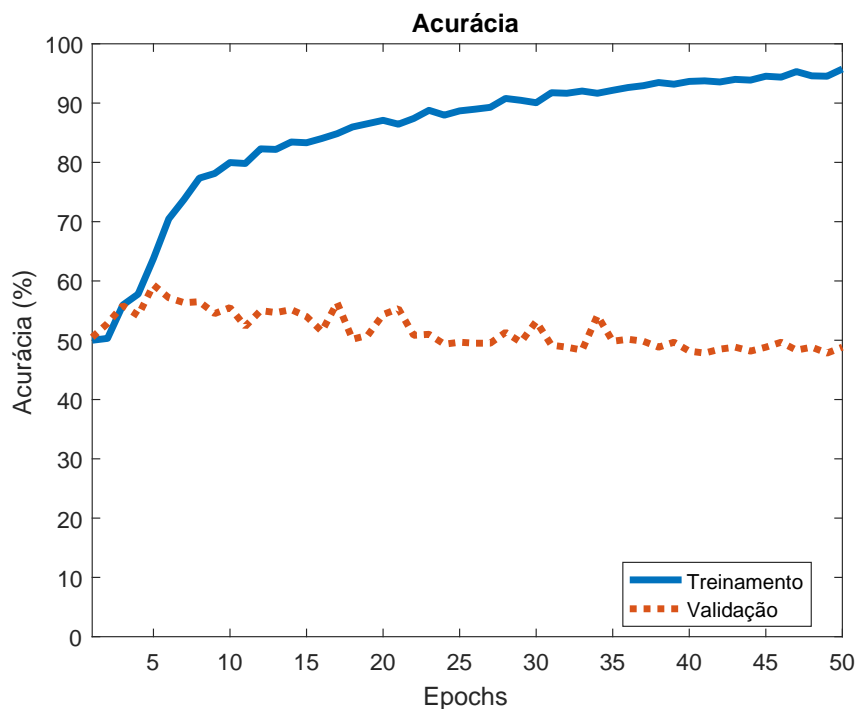
(c)

Figura 4.3: Classificação binária (Alzheimer x Saudáveis) no primeiro ambiente de experimentação - (b) Gráfico de perda x iterações ; e (c) Matriz de confusão.

## Classificação binária - Alzheimer x Distúrbios Cognitivos Leves

O segundo procedimento de classificação binária consiste em diferenciar portadores da doença e pacientes que apresentam distúrbios cognitivos leves, ou seja, pacientes que estão na progressão da doença. O total de imagens utilizada foi 4000, divididas na mesma proporção do cenário anterior. Podemos visualizar os resultados na Figura 4.4.

Analisando os gráficos de acurácia e perda, percebe-se que logo nas primeiras iterações o classificador já inviabiliza-se de distinguir as classes. A ocorrência de *overfitting* é destacada nesse cenário, evidenciando a semelhança entre as amostras.

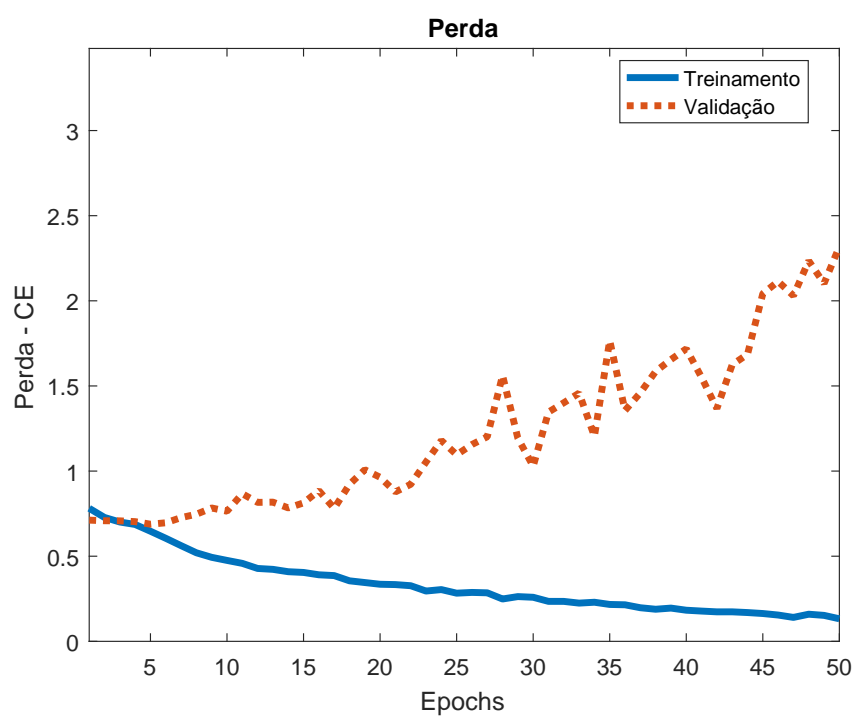


(a)

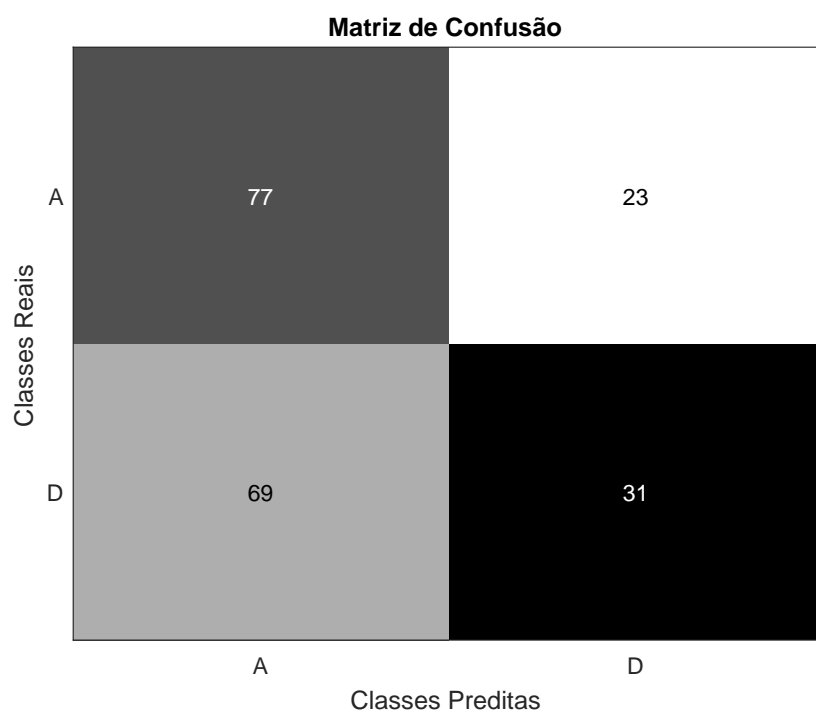
Figura 4.4: Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no primeiro ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Nesse cenário a dificuldade encontra-se em diferenciar classes que apresentam muitas semelhanças. Observa-se que a taxa de falsos positivos é elevada, demonstrando que as características dos portadores de distúrbios cognitivos leves favorecem à uma convergência errônea do classificador.





(b)



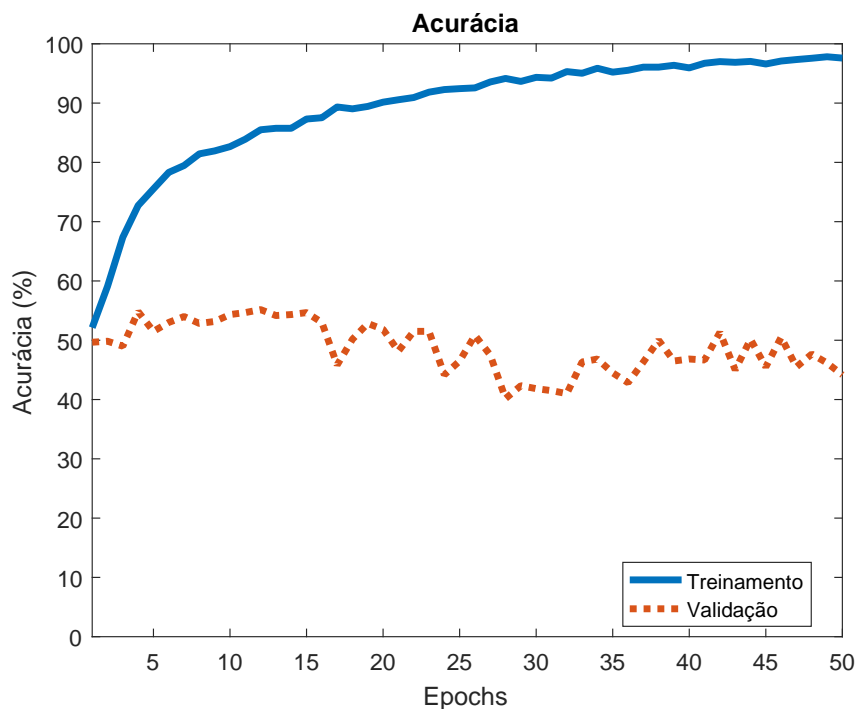
(c)

Figura 4.4: Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no primeiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

## Classificação binária - Distúrbios Cognitivos Leves x Saudável

O último cenário para o primeiro ambiente consiste em distinguir pacientes saudáveis de pacientes com distúrbios cognitivos leves. Nesse cenário a semelhança entre as amostras é maior, o que implica em resultados mais imprecisos do classificador. Podemos visualizar os resultados na Figura 4.5.

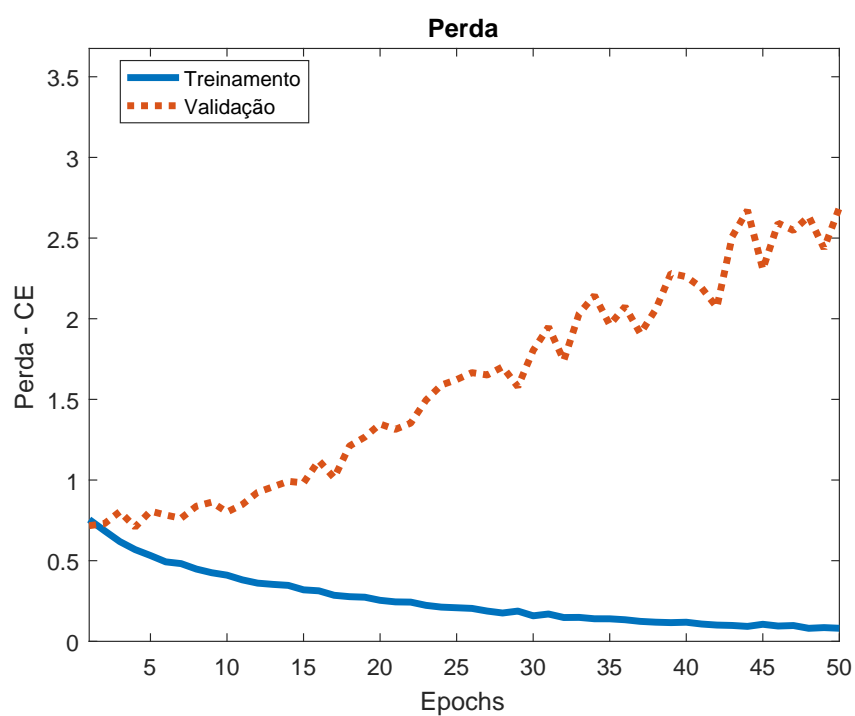
Como no cenário anterior, logo nas primeiras iterações, as curvas de acurácia e perda das amostras de validação divergem das curvas obtidas sobre o conjunto de teste, evidenciando a presença de *overfitting*.



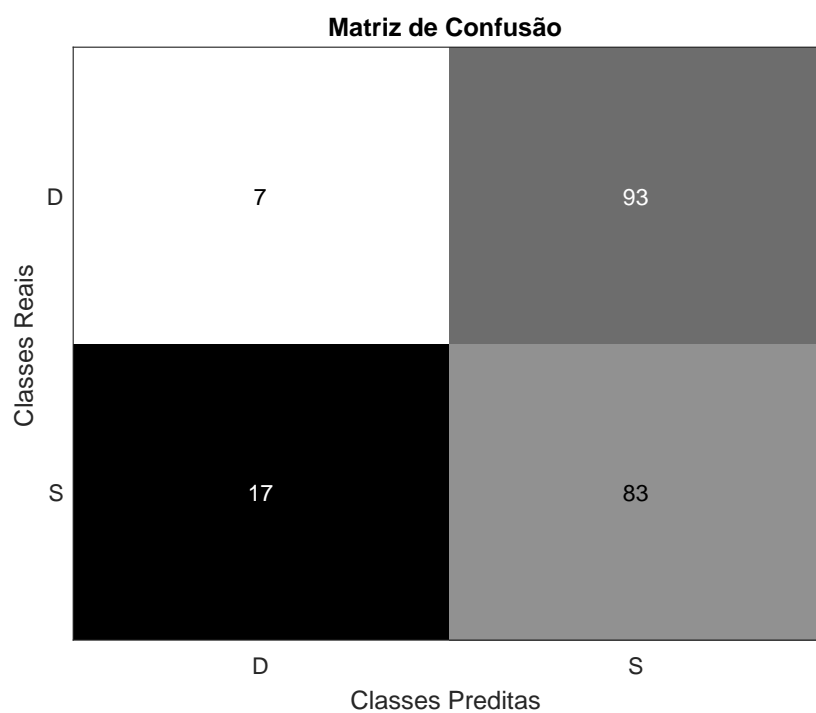
(a)

Figura 4.5: Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no primeiro ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Observando a matriz de confusão obtida, percebe-se uma alta taxa de falsos negativos, no qual podemos diagnosticar a maior semelhança entre amostras de pacientes saudáveis e pacientes com distúrbios, semelhança maior do que no cenário anterior entre pacientes com Alzheimer e pacientes com distúrbios.



(b)



(c)

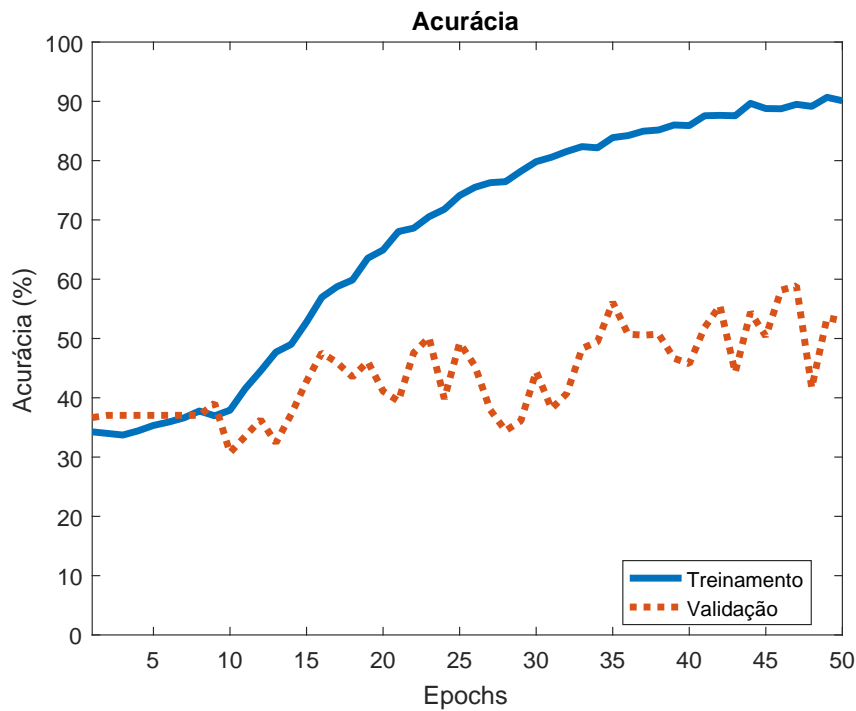
Figura 4.5: Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no primeiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

## 4.2.2 Segundo ambiente: conjunto reduzido

### Classificação multi classe

A classificação multi classe no segundo ambiente de experimentação consistiu em realizar a classificação das três classes, entretanto com um conjunto de amostras reduzido. Ao total, foram utilizadas 3660 imagens (1240 portadores da doença, 1240 com distúrbios cognitivos leves e 1180 saudáveis) divididas na seguinte proporção: 80% para treinamento, 15% para validação e 5% para teste. Os resultados podem ser visualizados na Figura 4.6.

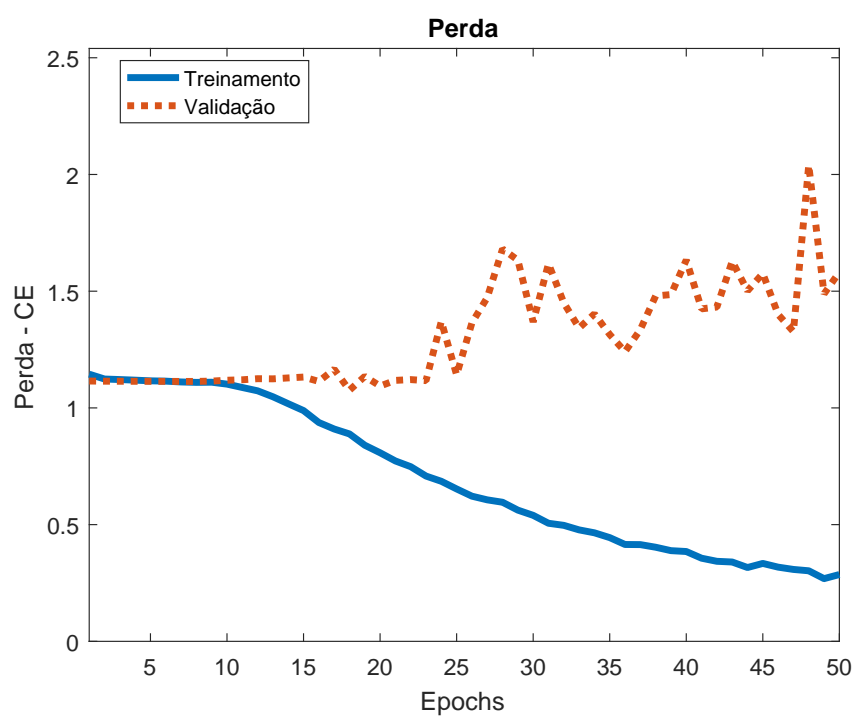
O gráfico de acurácia evidencia a imprecisão na classificação, uma vez que apresenta muita oscilação ao longo das iterações. Já o gráfico de perda apresenta uma curva ascendente sobre o conjunto de validação, ratificando a baixa acurácia.



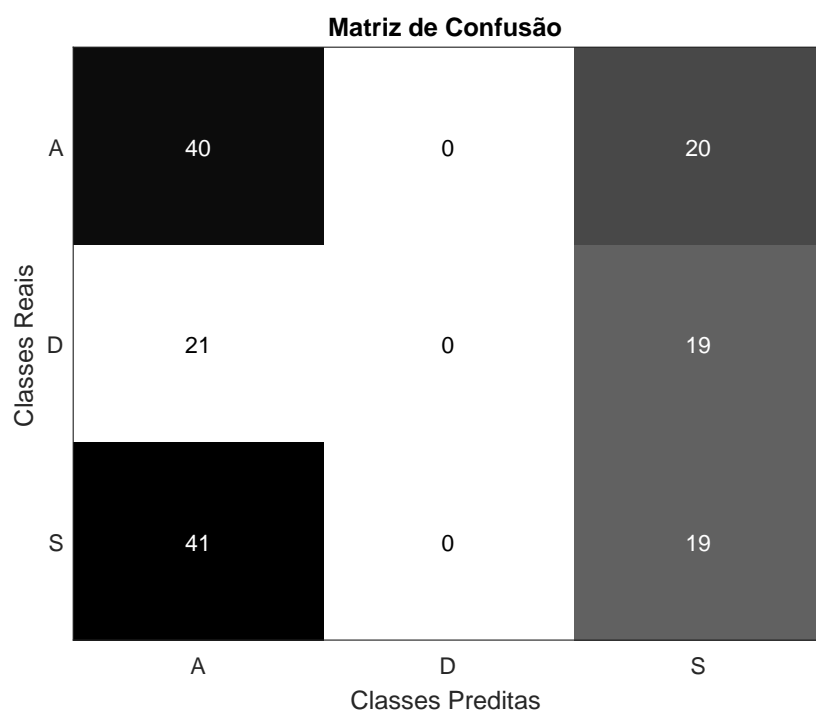
(a)

Figura 4.6: Classificação multi classe no segundo ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Nesse cenário, assim como no cenário multi classes do primeiro ambiente, o classificador não obteve parâmetros suficientes para diferenciar as três classes, designando a maioria das imagens como portadores da doença de Alzheimer e o restante como pacientes saudáveis.



(b)



(c)

Figura 4.6: Classificação multi classe no segundo ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

## Classificação binária - Alzheimer x Saudável

A primeira classificação binária no segundo ambiente comprovou que a baixa eficácia nesse cenário no primeiro ambiente era a semelhança entre as amostras. Após a remoção das amostras semelhantes, percebemos que as curvas de acurácia e perda apresentam comportamentos mais corretos (acompanhando a curva do conjunto de treinamento), mesmo com a oscilação da curva de perda para o conjunto de validação. Os resultados podem ser visualizados na Figura 4.7.

Nesse cenário foram utilizadas 2420 amostras (1180 de pacientes saudáveis e 1240 de portadores da doença), sendo divididas na seguinte proporção: 81% para treinamento, 14% para validação e 5% para teste.

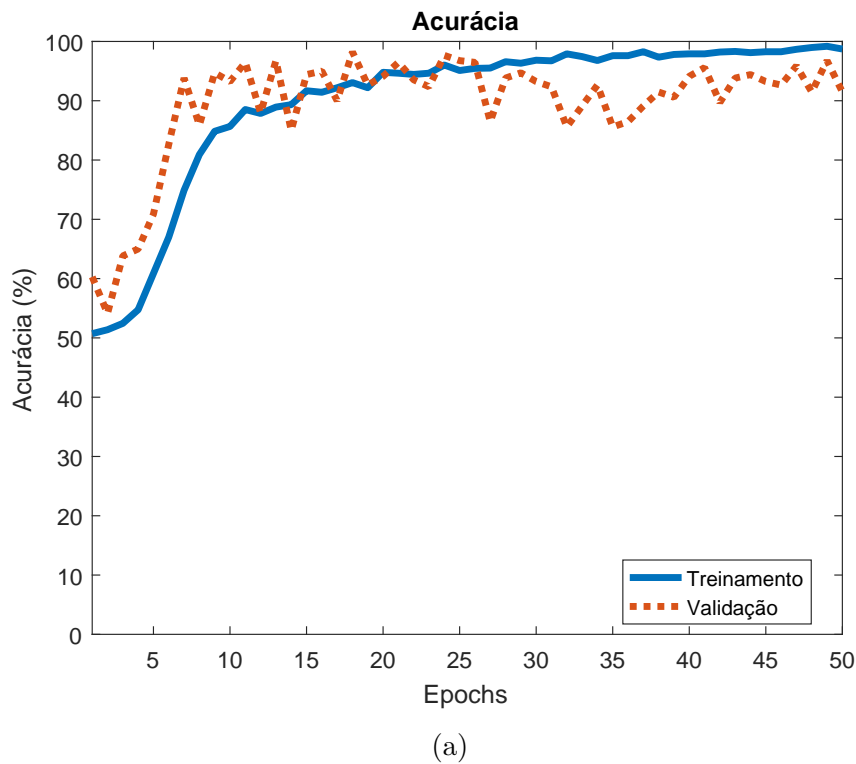
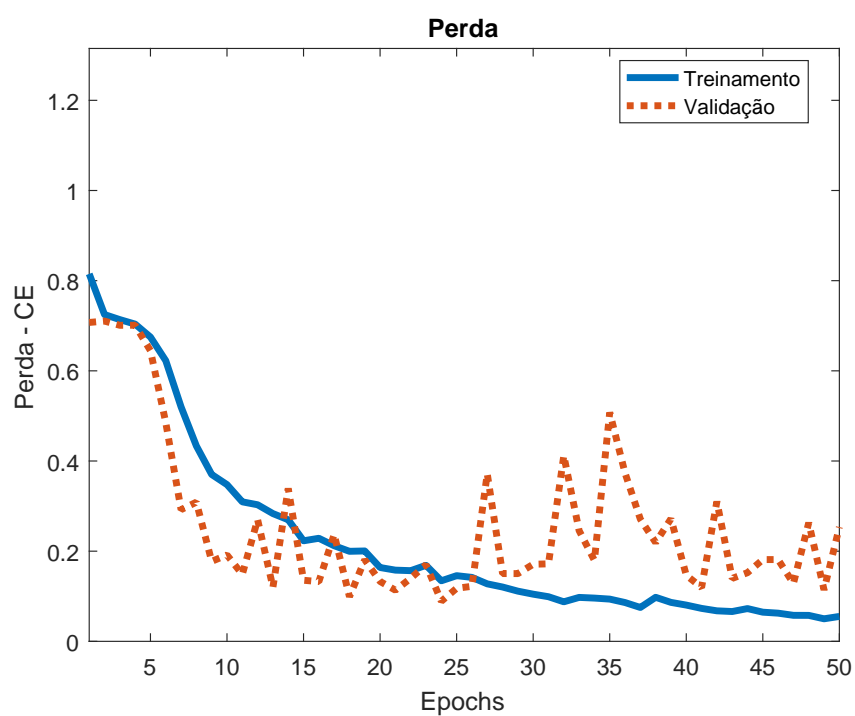
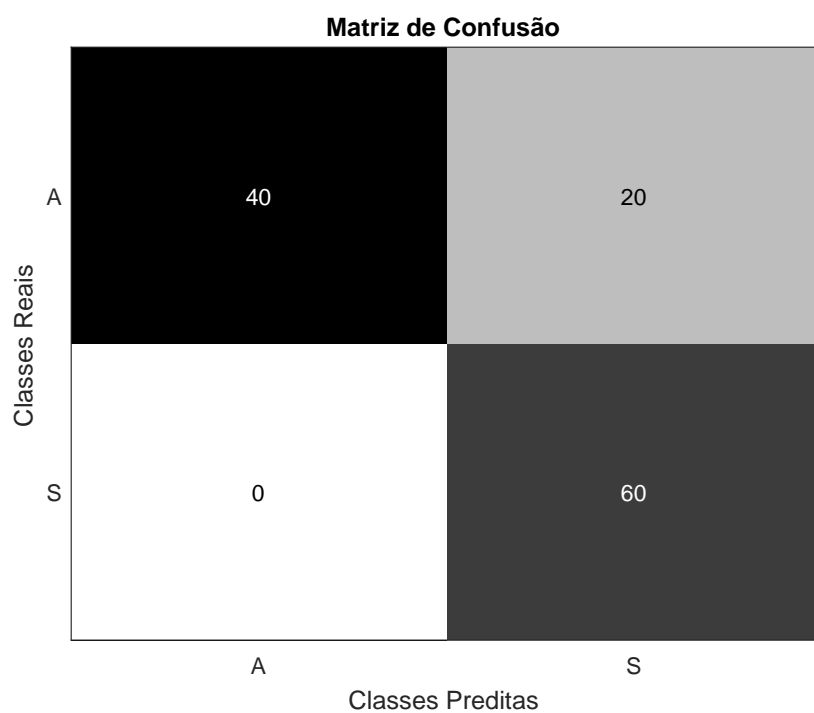


Figura 4.7: Classificação binária (Alzheimer x Saudáveis) no segundo ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Analisando a matriz de confusão percebemos a melhora efetiva nos resultados, classificando corretamente os pacientes saudáveis e errando em um terço das amostras dos portadores da doença de Alzheimer.



(b)



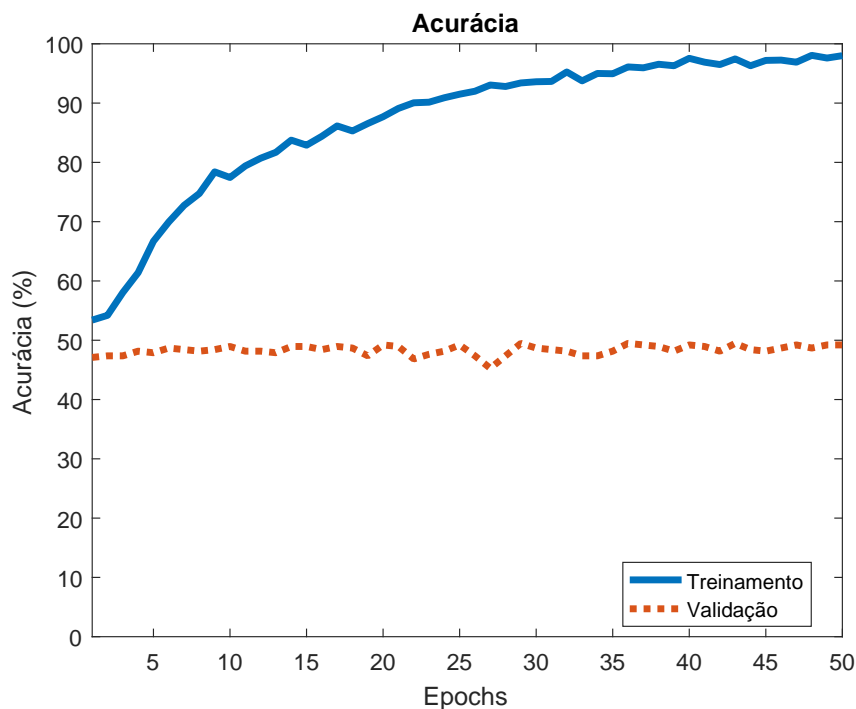
(c)

Figura 4.7: Classificação binária (Alzheimer x Saudáveis) no segundo ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

### Classificação binária - Alzheimer x Distúrbios Cognitivos Leves

Nesse cenário buscamos diferenciar os portadores da doença e os pacientes com distúrbios cognitivos leves. Foram utilizadas 2480 amostras (1240 de cada classe), sendo divididas da seguinte forma: 80% para treinamento, 15% para validação e 5% para teste. Podemos visualizar os resultados na Figura 4.8.

Analisando os gráficos de acurácia e perda, percebemos semelhança nas curvas obtidas nesse cenário e no cenário equivalente no primeiro ambiente de experimentação. Porém, mesmo com curvas semelhantes, os resultados foram melhorados.

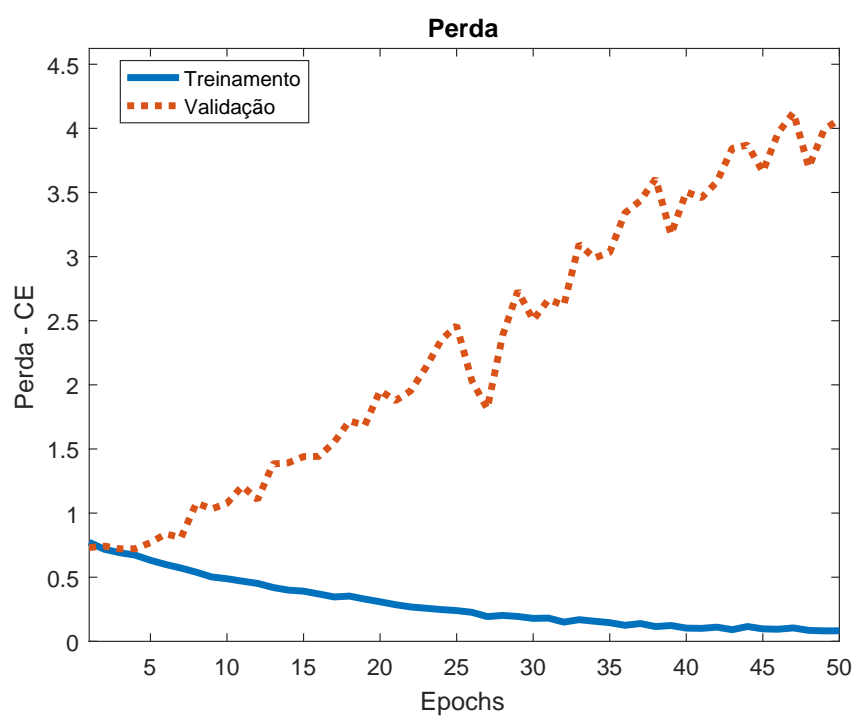


(a)

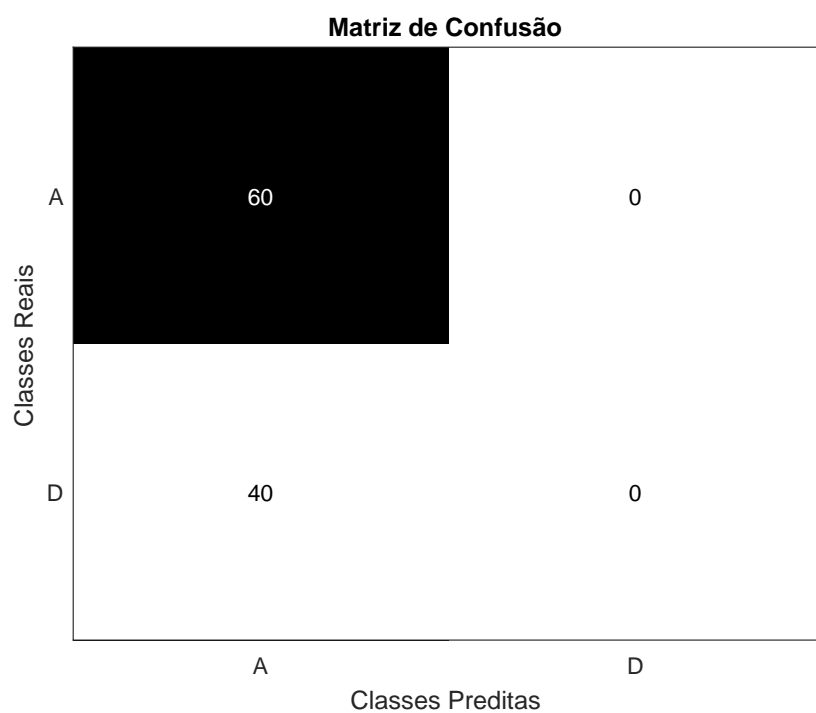
Figura 4.8: Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no segundo ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Observando a matriz de confusão percebemos que as amostras de pacientes com distúrbios cognitivos leves foram caracterizadas como portadores da doença, evidenciando que as características dessas amostras são de difícil diferenciação.





(b)



(c)

Figura 4.8: Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no segundo ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

### Classificação binária - Distúrbios Cognitivos Leves x Saudável

O último cenário para o segundo ambiente consiste em classificar os pacientes saudáveis e portadores de distúrbios cognitivos leves. Foram utilizadas 2420 amostras (1180 de pacientes saudáveis e 1240 de portadores de distúrbios cognitivos leves), sendo divididas na mesma proporção do cenário anterior. Podemos visualizar os resultados na Figura 4.9.

Observando os gráficos analisamos que as curvas de perda e acurácia, do conjunto de validação, apresentam comportamento próximos do ideal, acompanhando a curva dos dados de treinamento.

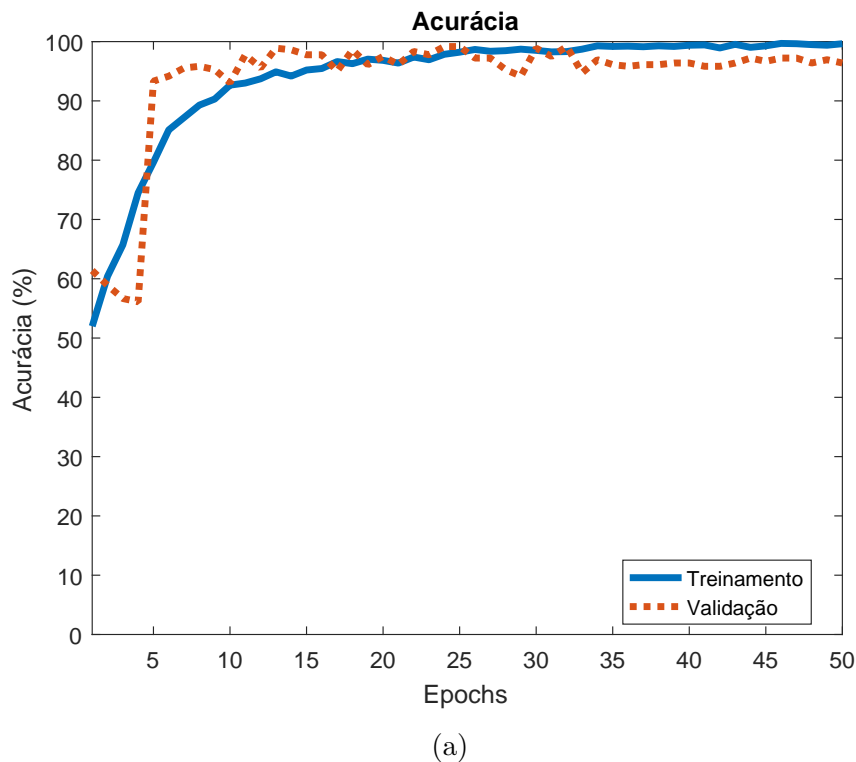
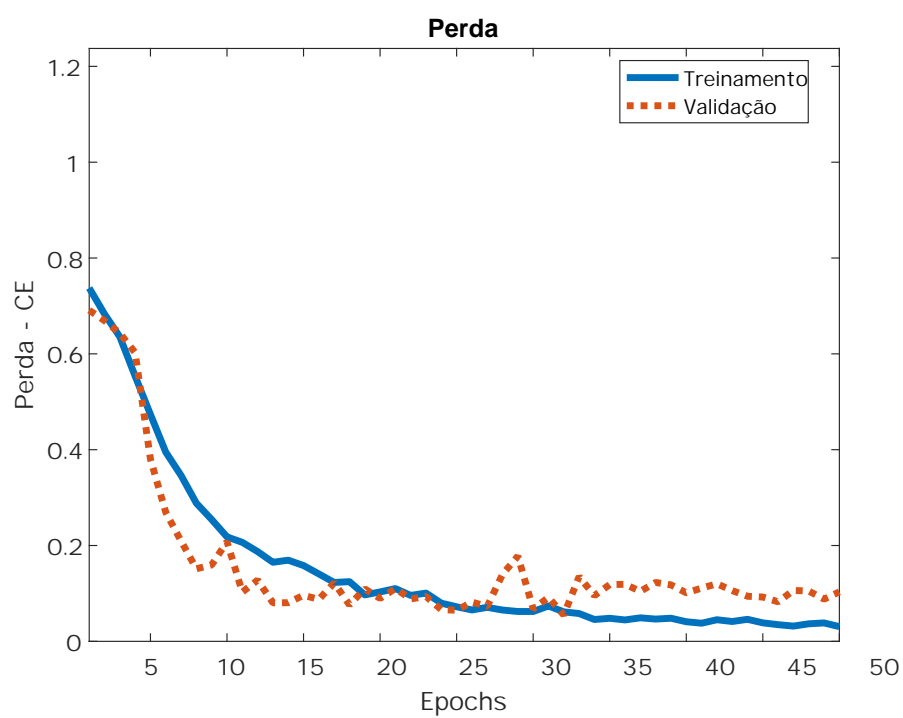
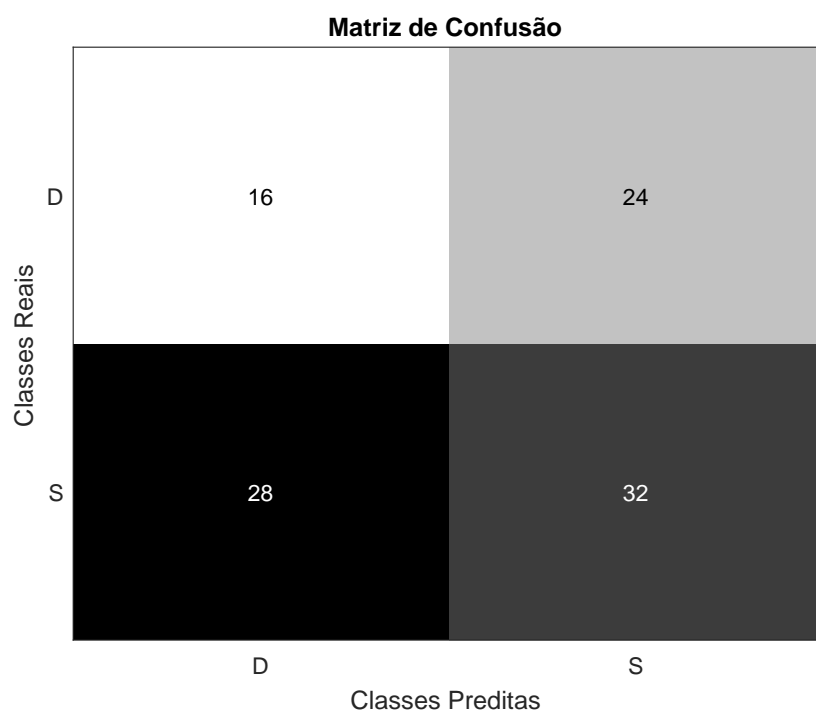


Figura 4.9: Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no segundo ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Mesmo obtendo gráficos com boa acurácia e baixa perda, a matriz de confusão demonstra resultados medianos, com muitos falsos positivos e muitos falsos negativos.



(b)



(c)

Figura 4.9: Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no segundo ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

### 4.2.3 Terceiro ambiente: conjunto reduzido com realce de bordas

#### Classificação multi classe

A classificação multi classe no terceiro e último ambiente de experimentação consistiu em realizar a classificação das três classes, entretanto com um conjunto de amostras reduzido e segmentado com o método *Canny*. Ao total, foram utilizadas 3660 imagens (1240 portadores da doença, 1240 com distúrbios cognitivos leves e 1180 saudáveis) divididas na seguinte proporção: 80% para treinamento, 15% para validação e 5% para teste. Os resultados podem ser visualizados na Figura 4.10.

Analisando o gráfico de perda percebemos que a curva das amostras de validação acompanha a curva das amostras de treinamento, entretanto esse acompanhamento é interrompido na iteração 10, o que ratifica a alta quantidade de amostras de pacientes com distúrbios cognitivos leves que foram classificadas de forma errada.

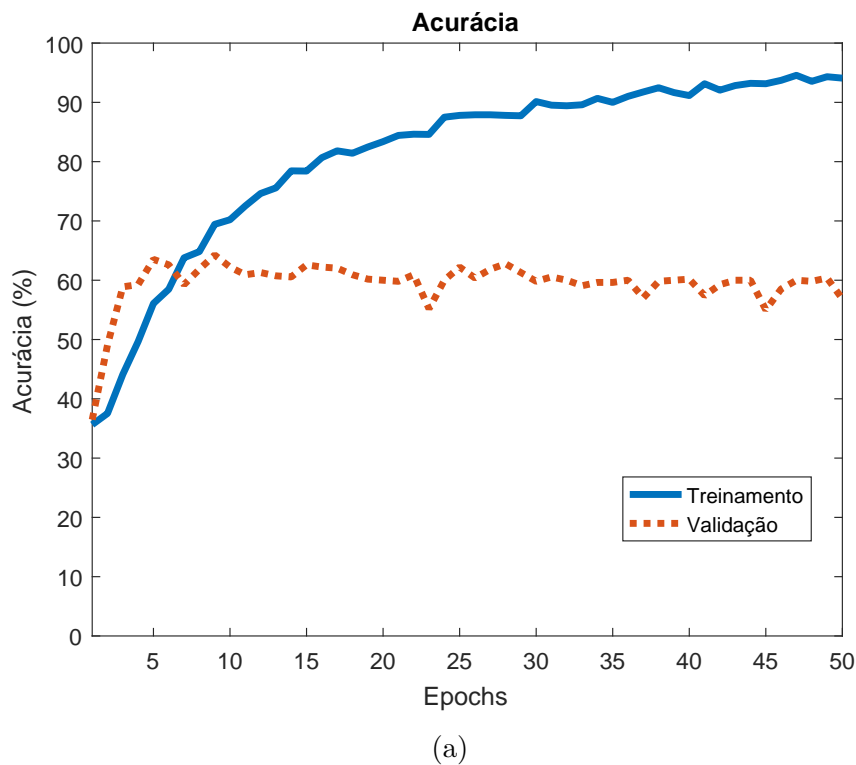
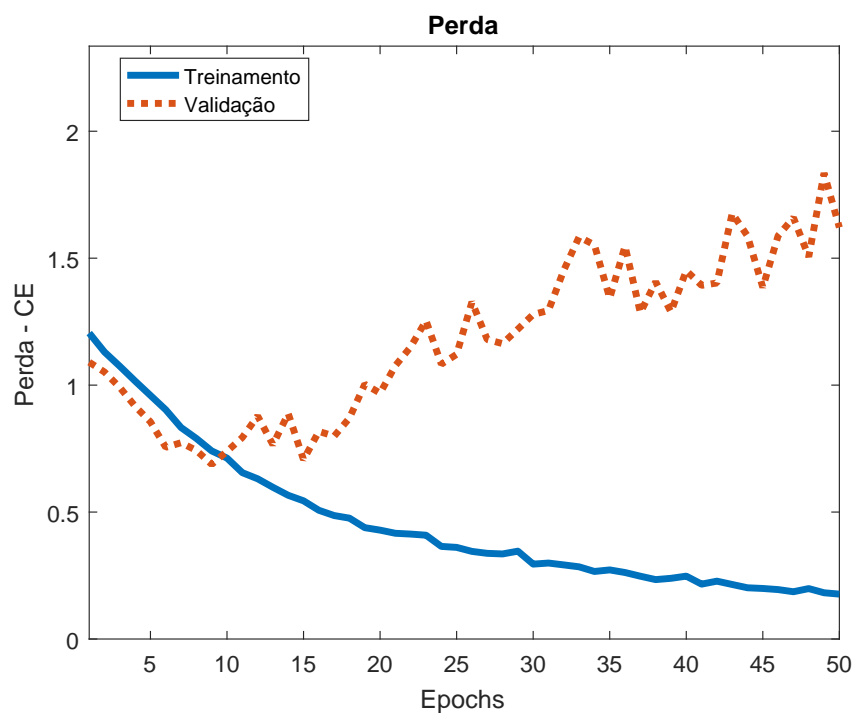


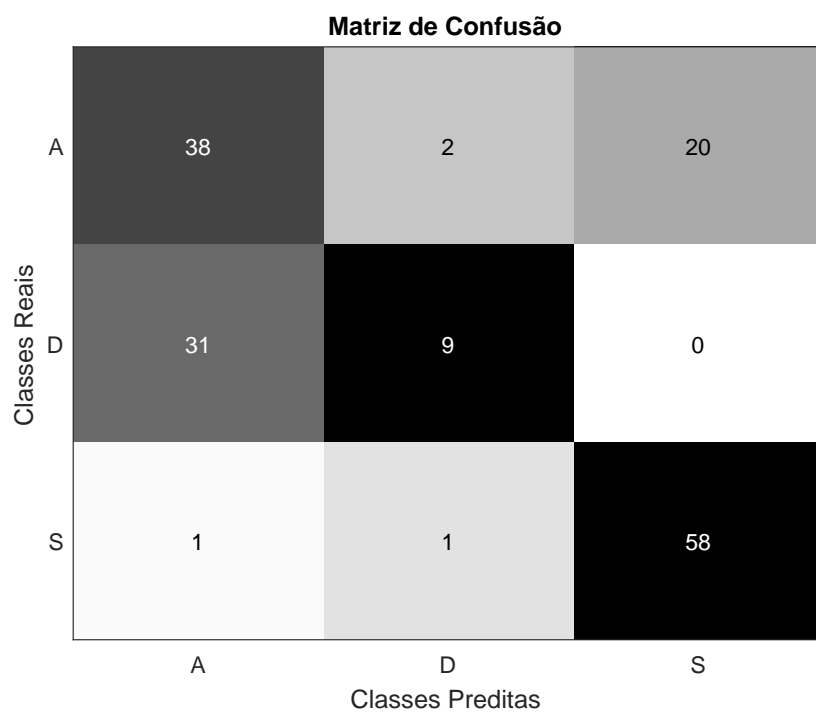
Figura 4.10: Classificação multi classe no terceiro ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Nesse cenário, percebemos um aumento considerável na performance do classificador, classificando de forma correta algumas amostras de pacientes com distúrbios cognitivos leves (situação que não ocorreu no respectivo cenário dos ambientes anteriores). Esse

aumento justifica-se pela presença mais nítida das regiões vazias nas imagens cerebrais dos pacientes.



(b)



(c)

Figura 4.10: Classificação multi classe no terceiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

### Classificação binária - Alzheimer x Saudável

A primeira classificação binária no terceiro ambiente, utilizando-se da segmentação, melhorou os gráficos de acurácia e perda obtidos no respectivo cenário no ambiente anterior. Os resultados podem ser visualizados na Figura 4.11.

Nesse cenário foram utilizadas 2420 amostras (1180 de pacientes saudáveis e 1240 de portadores da doença), sendo divididas na seguinte proporção: 81% para treinamento, 14% para validação e 5% para teste.

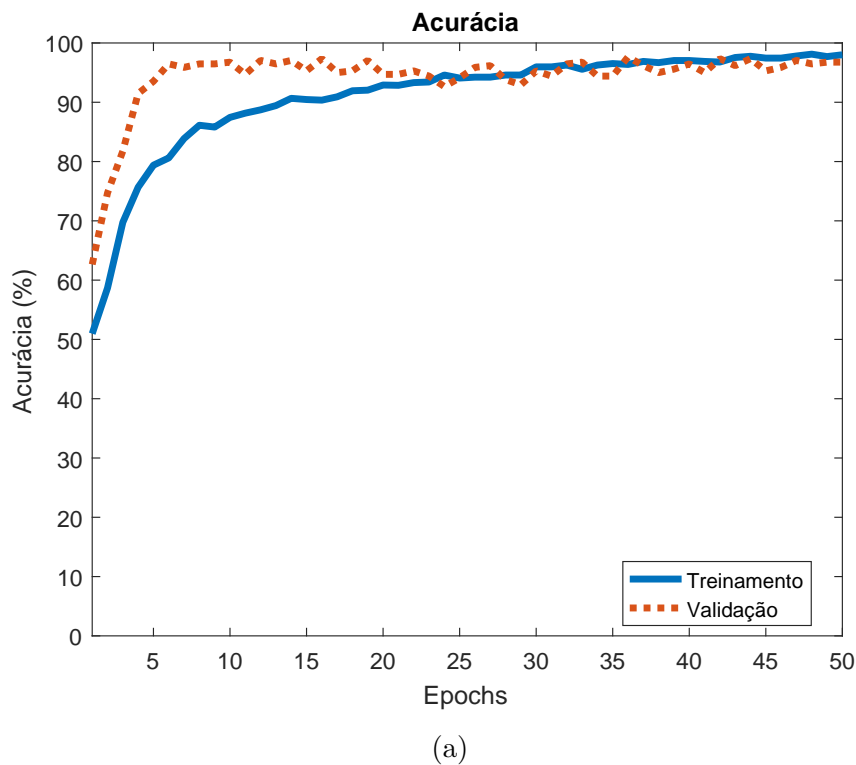
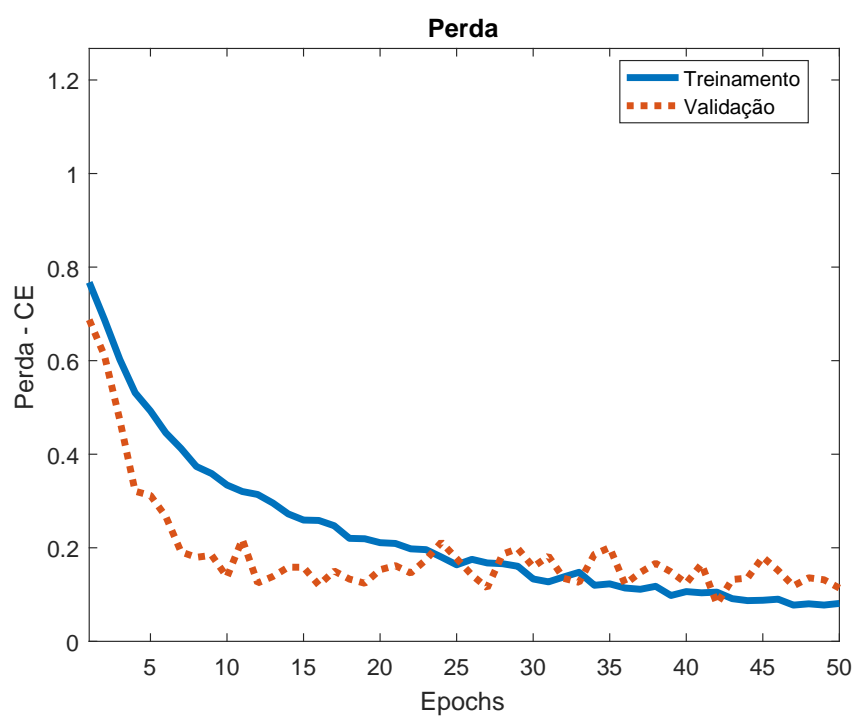
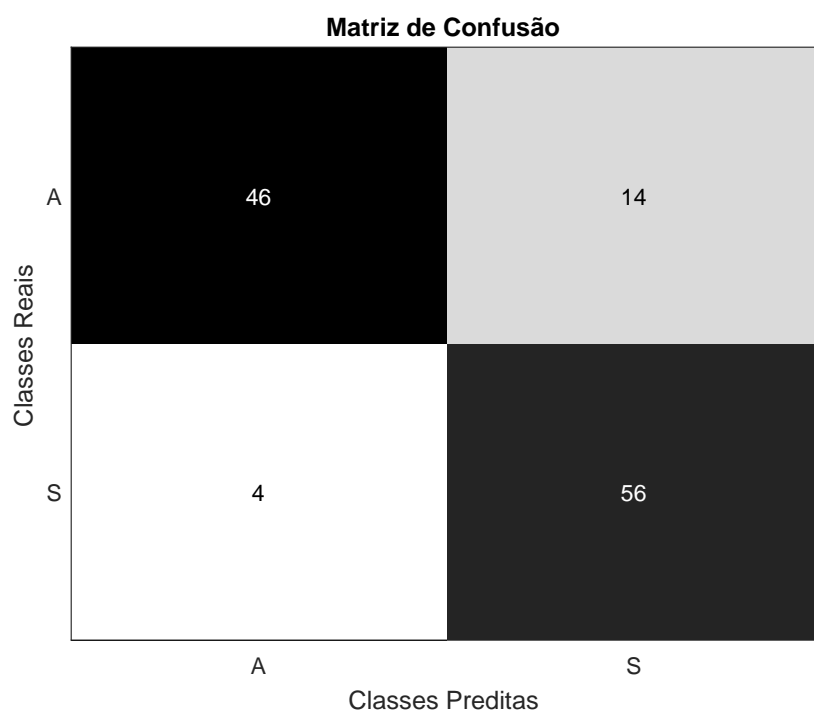


Figura 4.11: Classificação binária (Alzheimer x Saudáveis) no terceiro ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Ao contrário da matriz de confusão obtida nesse cenário no ambiente anterior, os resultados alcançados nesse ambiente foram mais equilibrados, com elevação da taxa de verdadeiros positivos.



(b)



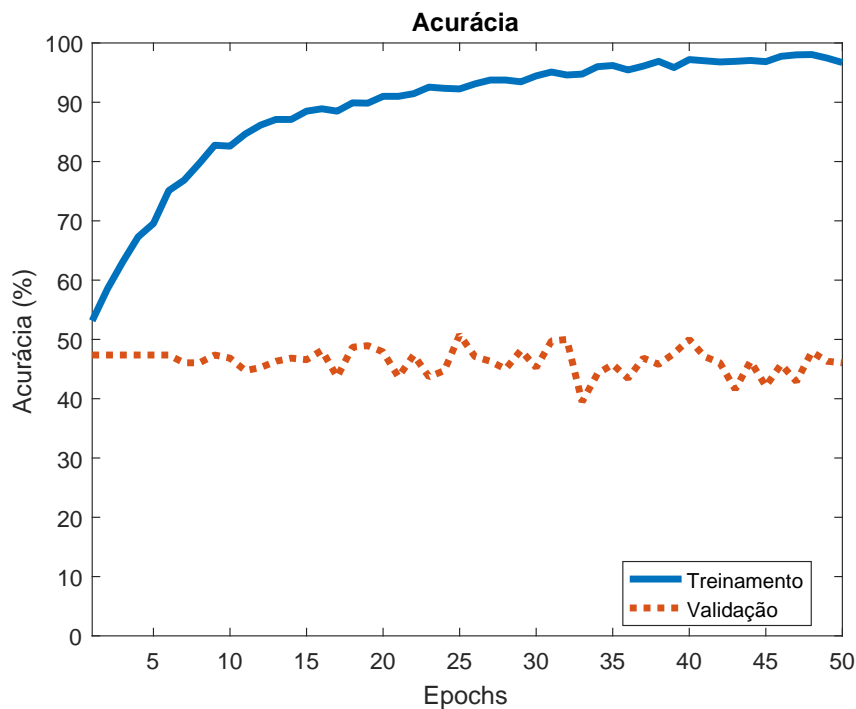
(c)

Figura 4.11: Classificação binária (Alzheimer x Saudáveis) no terceiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

### Classificação binária - Alzheimer x Distúrbios Cognitivos Leves

O segundo cenário de classificação binária consiste em avaliar melhorias ao respectivo cenário do ambiente anterior, dada a aplicação da segmentação. Foram utilizadas 2480 amostras (1240 de cada classe), sendo divididas da seguinte forma: 80% para treinamento, 15% para validação e 5% para teste. Podemos visualizar os resultados na Figura 4.12.

Analisando os gráficos de acurácia e perda, percebemos que foram produzidas curvas praticamente iguais às obtidas nesse cenário no ambiente anterior.

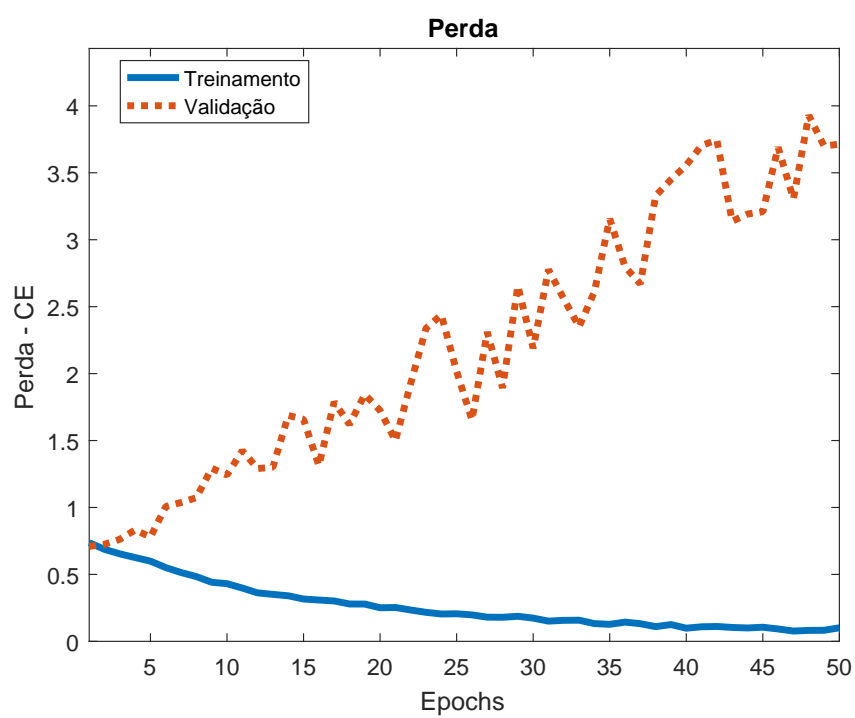


(a)

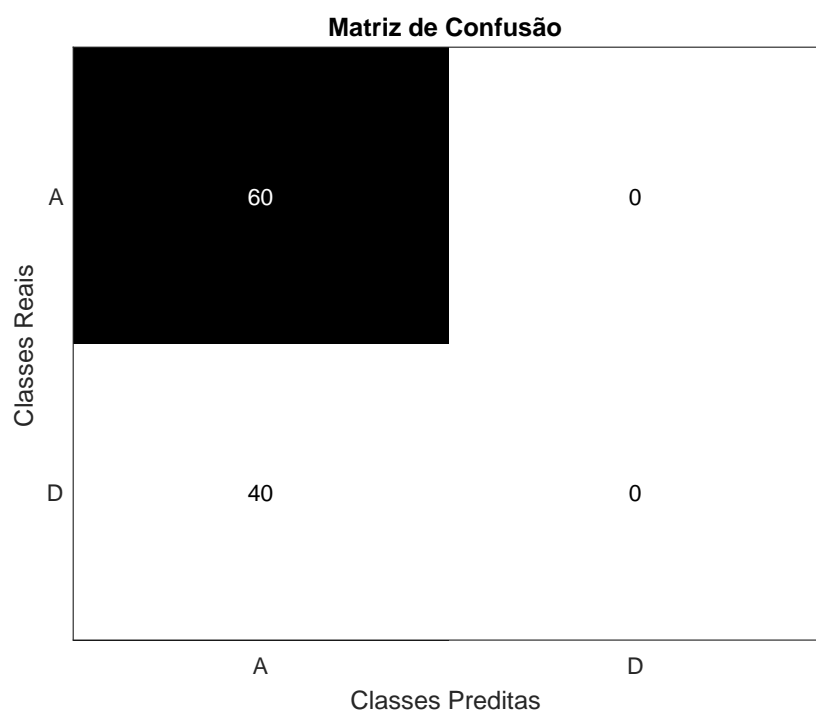
Figura 4.12: Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no terceiro ambiente de experimentação - (a) Gráfico de acurácia x iterações;

A proximidade entre as curvas obtidas, nesse e no ambiente anterior para o respectivo cenário, reflete-se em uma matriz de confusão idêntica. A classificação errônea de portadores de distúrbios cognitivos leves como portadores da doença ocorreu de forma similar.





(b)



(c)

Figura 4.12: Classificação binária (Alzheimer x Distúrbios Cognitivos Leves) no terceiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

### Classificação binária - Distúrbios Cognitivos Leves x Saudável

Por fim, o último cenário para o terceiro ambiente consiste em avaliar uma possível melhoria nos resultados obtidos no ambiente anterior, a partir da aplicação da segmentação. Foram utilizadas 2420 amostras (1180 de pacientes saudáveis e 1240 de portadores de distúrbios cognitivos leves), sendo divididas na mesma proporção do cenário anterior. Podemos visualizar os resultados na Figura 4.13.

Observando os gráficos analisamos que as curvas de perda e acurácia, do conjunto de validação, apresentam comportamento próximos do ideal, acompanhando a curva dos dados de treinamento.

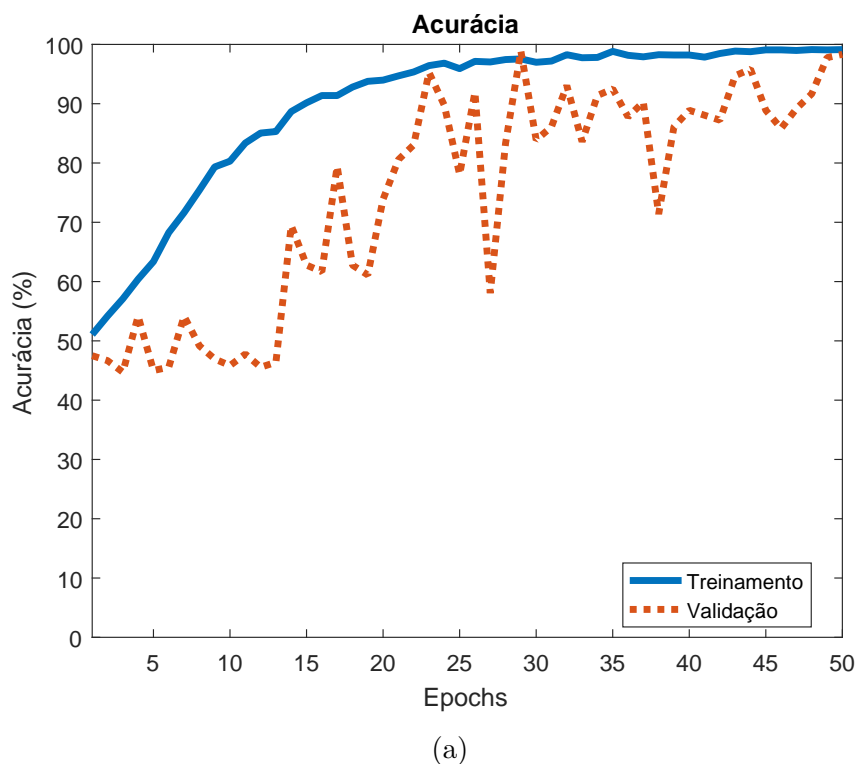
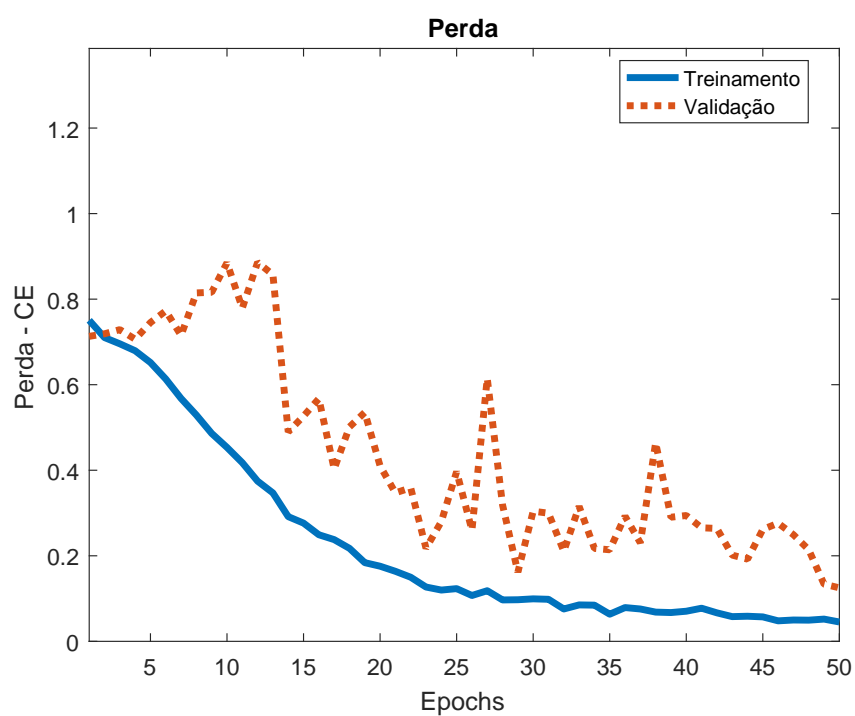
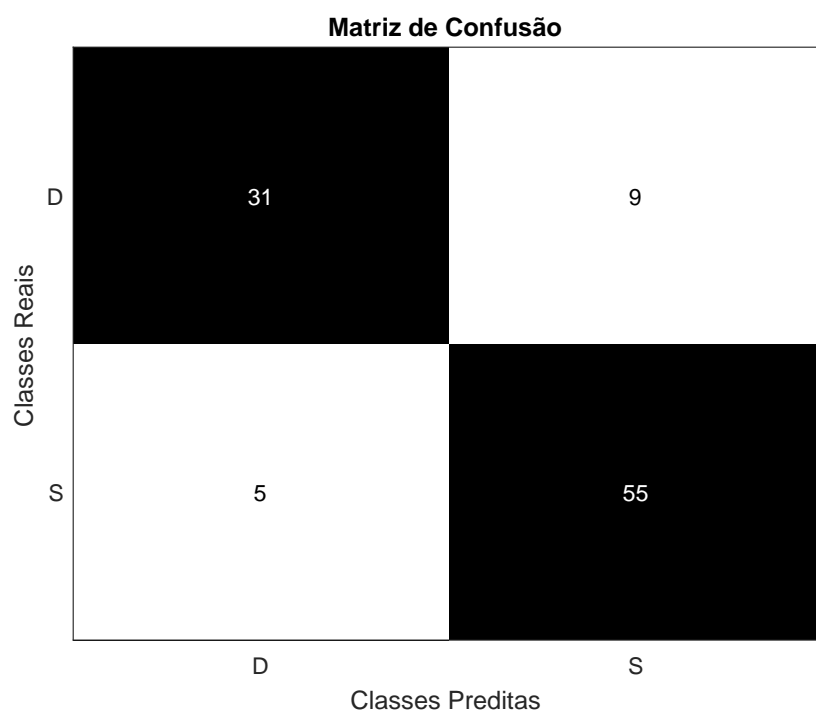


Figura 4.13: Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no terceiro ambiente de experimentação - (a) Gráfico de acurácia x iterações;

Apesar de obtermos gráficos de perda e acurácia com maiores oscilações, quando comparados aos obtidos no respectivo cenário no ambiente anterior, a matriz de confusão obtida para o conjunto de teste apresentou significativa melhoria, promovendo o alcance de resultados satisfatórios.



(b)



(c)

Figura 4.13: Classificação binária (Distúrbios Cognitivos Leves x Saudáveis) no terceiro ambiente de experimentação - (b) Gráfico de perda x iterações; e (c) Matriz de confusão.

# Capítulo 5

## Conclusão

A garantia de um diagnóstico preciso para inúmeras doenças passa pela capacidade técnica de um profissional da área da saúde, bem como da utilização de equipamentos de qualidade que auxiliem efetivamente nessa análise. Para a doença de Alzheimer, além da expertise, exames clínicos, dados demográficos e imagens médicas com qualidade razoável são fundamentais para percepção dos sintomas.

Entretanto, muitos profissionais, principalmente os que são iniciantes no ramo, carecem de expertise na análise das minúcias presentes nessas imagens, e que são cruciais para um diagnóstico confiável.

No momento do diagnóstico, o profissional precisa estar atento aos distúrbios motores e cognitivos, e que somente podem ser diagnosticados, na maioria das vezes, através da observação de regiões cerebrais responsáveis por essas funções. Essa análise é realizada sobre as imagens advindas de ressonância magnética, dependendo claramente da experiência técnica do profissional.

Obviamente essa expertise pode ser acumulada ao longo de sua trajetória profissional, porém, precisa ser sanada previamente, parcial ou totalmente, por alguma ferramenta, sendo o foco desse trabalho.

O trabalho aqui proposto busca realizar a análise de imagens médicas, através do processamento via redes neurais convolucionais que extraem características específicas que podem indicar ou não a presença da doença em um indivíduo.

O método elaborado aplica-se com melhor desempenho em cenários de classificação binária, com utilização de aprendizagem supervisionada. Observando os cenários executados, percebemos os melhores resultados em cenários onde os pacientes não estão em uma faixa intermediária da progressão da doença (distúrbios cognitivos leves). A aplicação da segmentação nas imagens mostrou um aprimoramento considerável nos resultados, removendo atributos desnecessários e realçando os relevantes.

Algumas melhorias possíveis para o trabalho em questão, sugeridas como trabalhos futuros, seriam a utilização das imagens digitais de diferentes eixos, sendo eles o sagital e o axial/transversal (não somente o coronal), implicando em um classificador bi ou tri dimensional; mesclar o classificador convolucional com uma rede neural convencional, que utilize dados clínicos, por exemplo, tipo sanguíneo, pressão arterial, temperatura corporal, etc., como entrada, e forneça um diagnóstico casado, gerando assim uma análise mais ampla do quadro patológico do paciente; e por fim variar a porcentagem de 55% de diferenciação entre as amostras ou criar um mecanismo de validação de amostras (amostras que não sejam de caráter ambíguo), utilizada nos ambientes 1 e 2, buscando perceber o impacto sobre o desempenho.

# Referências

- [1] Olivas, E. S., J. D. M. Guerrero, M. M. Sober, J. R. M. Benedito e A. J. S. Lopez: *Handbook Of Research On Machine Learning Applications and Trends: Algorithms, Methods and Techniques*, volume 1. IGI Global, 2009. x, 26
- [2] Nguyen, K., C. Fookes, A. Ross e S. Sridharan: *Iris recognition with off-the-shelf cnn features: A deep learning perspective*. Em *Visual Surveillance and Biometrics: Practices, Challenges, and Possibilities*. IEEE, 2017. x, 36, 37
- [3] Kumar, V., A. K. Abbas e J. C. Aster: *Bases Patológicas das Doenças*, volume 8. Elsevier, 2010. 1, 5, 6
- [4] Bear, M. F., B. W. Connors e M. A. Paradiso: *Neurociências - Desvendando o Sistema Nervos*, volume 3. Artmed, 2008. 3
- [5] Canny, J.: *A computational approach to edge detection*. Em *IEEE Transactions on Pattern Analysis and Machine Intelligence*, páginas 679 – 698. IEEE, 1986. 10
- [6] Slichter, C.: *Principles of Magnetic Resonance*, volume 3. Springer, 1992. 12
- [7] Larobina, M. e L. Murino: *Medical image file formats*. Em *Society for Imaging Informatics in Medicine 2013*, página 200–206, 2013. 12, 13
- [8] Kasabov, N. K.: *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering (Computational Intelligence)*, volume 1. The MIT Press, 1996. 13
- [9] Russel, S. e P. Norvig: *Artificial Intelligence: A Modern Approach*, volume 3. Pearson, 2009. 13
- [10] Fogel, D. B.: *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, volume 3. Wiley-IEEE Press, 2005. 13
- [11] Domingos, P.: *A few useful things to know about machine learning*. Commun. ACM, 55(10):78–87, outubro 2012, ISSN 0001-0782. <http://doi.acm.org/10.1145/2347736.2347755>. 14
- [12] Duda, R. O., P. E. Hart e D. G. Stork: *Pattern Classification*, volume 2. Wiley-Interscience, 2000. 15
- [13] Haykin, S.: *Neural Networks: A Comprehensive Foundation*, volume 2. Prentice Hall, 1998. 17

- [14] Aghdam, H. H. e E. J. Heravi: *Guide to Convolutional Neural Networks: A Practical Application to Traffic-Sign Detection and Classification*, volume 1. Springer, 2017. 23
- [15] Gunawardena, K A N N P, R. N. Rajapakse e N. D. Kodikara: *Applying convolutional neural networks for pre-detection of alzheimer's disease from structural mri data*. Em *2017 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*. IEEE, 2017. 27
- [16] Mayoraz, E. e E. Alpaydm: *Support vector machines for multi-class classification*. IDIAP, 1998. 28
- [17] Dumoulin, V. e F. Visin: *A guide to convolution arithmetic for deep learning*. ArXiv e-prints, março 2016. 28
- [18] Li, Fan, Danni Cheng e Manhua Liu: *Alzheimer's disease classification based on combination of multi-model convolutional networks*. Em *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*, número 61375112. IEEE, 2017, ISBN 9781538616208. 28
- [19] Nair, J. J. e N. Mohan: *Alzheimer ' s disease diagnosis in mr images using statistical methods*. Em *2017 International Conference on Communication and Signal Processing (ICCSPP)*, páginas 1232–1235. IEEE, 2017, ISBN 9781509038008. 29
- [20] Billones, C. D., O. J. L. D. Demetria, D. E. D. Hostallero e P. C. Naval: *Demnet: A convolutional neural network for the detection of alzheimer's disease and mild cognitive impairment*. Em *2016 IEEE Region 10 Conference (TENCON)*, páginas 3724–3727. IEEE, 2016, ISBN 978-1-5090-2597-8. 30
- [21] Simonyan, K. e A. Zisserman: *Very Deep Convolutional Networks for Large-Scale Image Recognition*. ArXiv e-prints, setembro 2014. 30
- [22] Cheng, D. e M. Liu: *Cnns based multi-modality classification for ad diagnosis*. Em *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, número 61375112. IEEE, 2017, ISBN 9781538619377. 30
- [23] Ollinger, J. M. e J. A. Fessler: *Positron-emission tomography*. Em *IEEE Signal Processing Magazine*, página 43–55, 1997. 30
- [24] Farooq, A., S. Anwar, M. Awais e S. Rehman: *A deep cnn based multi-class classification of alzheimer's disease using mri*. Em *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*, páginas 1–6. IEEE, 2017, ISBN 978-1-5386-1620-8. 31
- [25] Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke e A. Rabinovich: *Going Deeper with Convolutions*. ArXiv e-prints, setembro 2014. 31
- [26] He, K., X. Zhang, S. Ren e J. Sun: *Deep Residual Learning for Image Recognition*. ArXiv e-prints, dezembro 2015. 31

- [27] Shaha, M. e M. Pawar: *Transfer learning for image classification*. Em *Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA 2018)*. IEEE, 2018. 36
- [28] H., Wu e X. Gu: *Towards dropout training for convolutional neural networks*. CoRR, abs/1512.00242, 2015. <http://arxiv.org/abs/1512.00242>. 37
- [29] Laarhoven, T.: *L2 regularization versus batch and weight normalization*. 2017. 37
- [30] Patterson, J. e A. Gibson: *Deep Learning: A Practitioner's Approach*, volume 1. O'Reilly Media, 2017. 37
- [31] Sutskever, I., J. Martens, G. Dahl e G. Hinton: *On the importance of initialization and momentum in deep learning*. página 1139–1147, 2013. 38